



STATYSTYCZNIE RZECZ BIORĄC

JANINA

BAŃK

Czyli ile
trzeba zjeść
czekolady
żeby dostać
Nobla?

Janina Bąk

STATYSTYCZNIE RZECZ BIORĄC
Czyli ile trzeba zjeść czekolady, żeby dostać Nobla?



Dedykacja

Na początek: czy wiecie, że istnieje matematyczny wzór na dobre piwo?

Ilu osób potrzeba, by napisać jedną książkę?

Statystyka – królowa nauk czy popychadło medialnego świata?

Czy naukowcy są jak żony i nigdy się nie mylą?

Dlaczego koń i szop pracz nie dostały doktoratu?

Skąd wiemy, że owca rozpoznałaby na ulicy prezydenta?

Czy kiedy wychodzimy z psem na spacer, to średnio mamy po trzy nogi?

Czy 117% Polaków może się mylić?

Czy da się zmierzyć szczęście linijką?

Czy po seksie będziemy opanierowani w banknotach niczym krokiet w bułce?

Na imieniny lepiej kupić czekoladki czy kokainę?

Czy wszyscy jesteśmy rasistami?

Czy jak wsadzimy głowę do lodówki, a nogi do piekarnika, to będzie nam w sam raz?

Ile trzeba zjeść czekolady, żeby dostać Nobla?

Czy język angielski powoduje zawały serca, a masło przedłuża życie?

Czego statystyka nauczyła się od martwego łososia?

Dlaczego ufamy szarlatanom?

Na koniec: dlaczego wciąż się wzruszam na myśl o bezdomnych
lamach?

Bibliografia

Przypisy

Krzysiowi, który nauczył mnie, że bycie bliźniakiem to nie tylko urodziny tego samego dnia, a potem milion razy udowadniał, że to prawda. A ja studiowałam statystykę i wiem, że milion to dużo.

Na początek: czy wiecie, że istnieje matematyczny wzór na dobre piwo?

Jeden z najbardziej znanych statystyków na świecie wcale nie był człowiekiem-akademikiem. Nie przychodził do roboty w piżamie, nie pisał w uczelnianej windzie markerem rozszerzenia liczby pi, a na śniadanie nie jadał kanapek z artykułami naukowymi zamiast sałaty. Być może trudno w to uwierzyć, ale jego życie było znacznie bardziej ekscytujące. Otóż ów człowiek zawodowo warzył piwo, i to jedno z najlepszych piw na świecie – irlandzkiego Guinnessa. Nazywał się William S. Gosset¹, a oprócz doskonałego piwa zawdzięczamy mu też jeden z najpopularniejszych testów statystycznych, test t-Studenta², i pierwsze rozważania dotyczące istotności statystycznej. Cel tych badań był szczytny – opracowanie matematycznej metody, która gwarantowałaby, że absolutnie każde piwo opuszczające browar Guinness będzie tak samo dobre jak wzorcowe. Czyli to w sumie dokładnie to, czym zajmują się absolutnie wszyscy studenci w krajowych akademikach, z tym że oni poszli raczej w zaawansowane badania empiryczne. Klawo, nie? Statystyka pozwala stworzyć matematyczny wzór na dobre piwo! Zresztą nie tylko. Również wzór na miłość (Skopek et al. 2010), a także – o czym piszę w rozdziale *Czy da się zmierzyć szczęście linijką?* – na radość życia. I właśnie o tym jest ta książka. O tym, do czego właściwie przydaje nam się statystyka, i – że do wszystkiego.

Dlaczego wyniki badań są lepsze od obserwacji pani Hani spod piątki, która przecież od trzydziestu lat siedzi w oknie i wie wszystko o życiu osiedla?

Czy rację miał Kazimierz z Białej Podlaskiej, oburzony tym, że ta ankieta, o której przeczytał w gazecie, nadaje się tylko do kosza, bo podobno przeprowadzono ją wśród bałaganiarzy, a jego nikt o zdanie nie spytał, mimo że od 1995 roku nie może znaleźć swojego chomika, który niefortunnie zawieruszył się wśród brudnych naczyń?

Dlaczego twierdzimy, że palenie powoduje raka płuc, ale gdy babcia mówi, że podczas ataku halnego jej sernik, a w konsekwencji i ręce, zawsze opadają, to komitet noblowski nie jest skłonny, by uznać, że jedno

powoduje drugie? No więc w jaki sposób ustalić te związki przyczynowo-skutkowe?

Dlaczego istnieją sytuacje, w których świat nauki nie rzuca jajkami w autorów badań, chociaż przebadali ledwie kilka osób? Czy to prawda, że wielkość (próby badawczej) nie ma znaczenia?

Jak to możliwe, że przez wszystkie lata szkolne liczono nam średnią ocen i to było OK (choć nasi rodzice niekoniecznie twierdzili, że OK), a jednak zawsze znajdzie się jakiś fanatyk (statystyk), który uważa, że kiedy jesteśmy z psem na spacerze, to statystycznie wcale nie mamy trzech nóg?

Czy możemy ufać tym miłym naszemu sercu gazetowym nagłówkom, które donoszą: „uprawiaj więcej seksu, a zarobisz więcej”? (Spoiler: nie. Ale zawsze można próbować). W jaki sposób zweryfikować, czy odkrycia amerykańskich naukowców mają sens, czy też wcale nie jest prawdą, że – jak donosi portal Fronda – omlet z pieczarkami zwalczy depresję?

Nie można pominąć też innych, kluczowych dla rozwoju ludzkości zagadnień. Czy na imieniny lepiej kupić czekoladę, czy kokainę? Skąd wiemy, że owca rozpoznałaby na ulicy Zenka Martyniuka? Czy wielkość penisa wpływa na rozwój gospodarczy kraju? To tylko kilka z wielu pytań, na które odpowiada ta książka.

Nie czytajcie jej w kapciach, bo wam pospadają z wrażenia.

Ilu osób potrzeba, by napisać jedną książkę?

To nie jest podręcznik do statystyki, nie ma więc potrzeby dzwonić do zakładu pogrzebowego, że halo, proszę przyjechać, zaraz umrę z nudów. Bo to książka o tym, że statystyka jest wspaniała. Lubię myśleć o tej dyscyplinie jak o takim pomelo naukowego świata – mnie wiecznie zadziwia, wielu onieśmiela. No bo, dajcie spokój z tym pomelo, przecież człowiek nigdy nie wie, czy to grejpfrut jest, czy pomarańcza, czy też owoc (he, he) nieudanego romansu cytrusa z arbuzem. W sumie to nie wiadomo³. No właśnie, to jest to, czego nauczyłam się po latach zapasów z tą dyscypliną – że niczego nie można być pewnym. A najbardziej fascynujące w tym wszystkim jest to, że statystyka potrafi tę niepewność policzyć. Pozwala również oszacować matematyczny wzór na miłość. I w ogóle rozplata misterną sieć związków przyczynowo-skutkowych we wszystkim, co nas otacza.

W życiu zawodowym pozeniłam statystykę z naukami społecznymi, głównie socjologią migracji i socjologią pracy. Daleko mi do wielkich matematyków wymyślających zmieniające świat wzory, często niedostępne dla zwykłych śmiertelników – mój podziw dla takiej pracy jest nieskończony, ale ja poszłam w trochę inną stronę. Tworzyłam modele pozwalające przewidzieć, kto w ciągu kolejnych kilku lat awansuje w strukturach danej organizacji. Pisałam równania, które miały wyłuskać wszelkie nierówności rządzące rynkiem pracy. Projektowałam eksperymenty, których celem było wyjaśnienie, dlaczego ludzie decydują się przyjąć propozycję jednego pracodawcy, a zrezygnować z innej. Pracowałam z liczbami, modelami i równaniami, którymi próbowałam objaśnić wszelkie zagadki społecznego świata, wypełniałam puste pola wiedzy o człowieku, rozplatałam skomplikowane struktury ekonomii behawioralnej, a potem przychodziłam na zajęcia i student pytał mnie, czy da się oszacować, jakie jest prawdopodobieństwo, że w swoim życiu spotka w lesie dwa razy tego samego jeża.

Oczywiście, że się da⁴. Dajcie spokój, ta dyscyplina nie miałaby żadnego sensu, gdyby nie umiała dawać odpowiedzi na tak fundamentalne pytania. Albo szerzej – zaspokajać ludzkiej ciekawości. Jestem wdzięczna, że przez

kilka lat mogłam towarzyszyć moim studentom w tych dociekaniach i związanej z nimi niepewności. Przecież nie oszukujmy się, to nie zawsze jest prosta dyscyplina. Mam jednak nadzieję, że zapamiętali z tych zajęć to, co chciałabym, byście i wy zapamiętali z tej książki – że statystyka to nie tylko wzory, trudne definicje i dziwne znaczki, których nie sposób znaleźć na klawiaturze i potem człowiek, kiedy pisze pracę magisterską, cały czas musi kopiować grecką betę z Wikipedii. Statystyka to wspaniałe narzędzie do odpowiadania na wszelkie pytania dotyczące otaczającego nas świata, na przykład – choć nie tylko – na te dotyczące jeży. Jest to też dyscyplina dość smutna w tym sensie, że gdyby była psem, to najpewniej takim łaciatym kundelkiem o wielkości konia, którego wszyscy się boją, a który po prostu bardzo chce być kochany. Nie no, serio, nie bójcie się, po prostu spróbujcie go oswoić.

Myślę, że to, co jest najtrudniejsze w uczeniu się statystyki, to głębokie przekonanie wielu ludzi, że kompletnie się do tego nie nadają. Najczęściej dlatego, że ktoś im kiedyś powiedział, że nie dadzą rady, że się tego nigdy nie nauczą, że to nie dla nich. Lubię powtarzać, że nie ma głabów matematycznych, są tylko źli nauczyciele. Że każdy może się nauczyć poprawnego wnioskowania i mentalnie poprzytulać się z dowolnym modelem regresji. Jeśli w jakiejś szufladce w głowie trzymacie kartkę, na której ktoś kiedyś napisał wam: „humanistyczny umysł”, „głęb matematyczny”, „liczbowy analfabeta”, to jest to najlepszy moment, by ją zgnieść i rzucić na pożarcie kotom. Kto nie ma kota w domu, może ją wyrzucić do śmieci. Byłe do segregowanych.

Gdy będziecie czytać tę książkę i nieśmiało pluskać się w jeziorcu metodologicznych pojęć i sadzawce metody naukowej, to chciałabym, byście myśleli sobie o... dzieciach. Popatrzcie na to w ten sposób: nikt nie śmieje się z niemowlaków, że nie potrafią mówić (a jeśli tak, to tylko za ich plecami i na tyle cicho, żeby nie słyszały). Jest dla nas naturalne, że to umiejętność, której dopiero muszą się nauczyć, jeszcze trochę poczekać i poćwiczyć.

Tak samo jest ze statystyką. Jednak tu przy pierwszych trudnościach natychmiast stwierdzamy, że dość tych matematycznych upokorzeń. Oto najwyższy czas zamknąć się w piwnicy u rodziców i do końca życia wycinać anonimy z literek z gazety, z prostym, acz wyrazistym przesłaniem: **PRECZ Z LICZBAMI!!!**

Mam w sobie dużo miłości do liczb, statystyki i metodologii. I to się wspaniale składa, bo gdyby nie miłość, nie byłoby tej książki. Dziękuję za to, że powstała, trzem mężczyznom, którzy wierzyli we mnie bardziej niż ja sama.

MOJEMU MĘŻOWI, najdoskonalszemu współczynnikiowi we wszelkich moich życiowych równaniach. Jest bozonem Higgsa mojego wszechświata – bez niego nie byłoby niczego.

MAĆKOWI MAKSELONOWI, bo choć wedle umowy miał być taką moją redaktorską mamą kwoką, to na końcu procesu więcej jednak było w tym wszystkim przyjaźni i nieskończonych pokładów cierpliwości. Na przykład wtedy, gdy zapytał mnie, kiedy w końcu napiszę rozdział, a ja musiałam mu powiedzieć, że chwilowo nie mogę, bo niosę przez miasto wiklinowego kangura i bardzo jest nieporęczny. I Maciek mnie nie zabił. A mógł.

PIOTROWI BUCKIEMU, który jest człowiekiem o tak pięknym umyśle i pluszowym sercu, że każda rozmowa z nim była jednocześnie jak naturalna tabletką uspokajającą na niepewność lub – gdy była odwrotna potrzeba – jak mocne intelektualne espresso (z tym że bez cierpkiego posmaku, sama słodycz).

Dziękuję prawdopodobieństwu, które nas sobie przytrafiło.

Gdyby nie oni, nie powstałaby ta książka. A gdyby nie powstała, to nie miałabym możliwości przekonania was, że to brzydkie kaczątko wśród dyscyplin naukowych jest warte uwagi. Dla większości ludzi statystyka to taki brzydki, smutny chłopak w odziedziczonych po dziadku sztruksowych spodniach, który przyszedł na potańcówkę, bo matka mu obiecała trzy złote i chupachupsa, jeśli pójdzie (bo w końcu kiedyś musi wyprać tę jego pościel z Supermanem), i on przywdział te sztruksy, marynarkę z łatami na łokciach i jeszcze, by dodać całości nonszalancji, T-shirt z napisem FBI. FEMALE BODY INSPECTOR (którego znaczenia nie rozumie, bo w jego wsi nieprzerwanie od 1772 uczą rosyjskiego), pożegnał czule swoją dziewczynę, co ją sobie jakiś czas temu wyciął z kartonu po mleku, żeby rodzina przestała się czepiać, skąpał się w delikatnej chmurze hjugobosa kupionego od luja spod bloku za konia i słoik kapuśniaku matki, i poszedł na tę potańcówkę, a potem osiem godzin stał samotnie pod ścianą. Nikt z nim nie chciał zatańczyć, nikt porozmawiać, królem balu też nie został. Jeśli więc statystyka jest tym biednym chłopcem w sztruksach na wiejskiej potańcówce, to chciałabym, by ta książka pełniła funkcję opisu na Tinderze. Ale wiecie, takiego, który zwala z nóg i sprawia, że w pierwszej

wiadomości pytacie: „Cześć, czy macie czas wieczorem i ile dzieci mogą Ci urodzić?”.

I nie, książka też nie odpisuje na takie wiadomości. Ale nauczy was przynajmniej te swoje dzieci liczyć.

Statystyka – królowa nauk czy popychadło medialnego świata?

W 2001 roku na uniwersytecie w Wirginii profesor Timothy Wilson przeprowadził eksperyment. Wymyślił sobie, że będzie razić ochotników prądem. Wiadomo – nauka. I dostał na to grant, o czym wam mówię, gdybyście się kiedyś zastanawiali, na co idą wasze podatki. Na potrzeby eksperymentu skonstruował specjalny przycisk, który mieli naciskać badani. I wtedy szok (dosłowny, he, he, he) – przyciskający byli rażeni prądem. Ale nie takim byle jakim prądem, co to człowiekowi po palcach przebiega, jak mu się kot naelektryzuje balonem – to był taki fizyczny, realny ból o bardzo wysokim natężeniu. Taki ból, jaki odczuwa się w trakcie oglądania polskich komedii romantycznych. Następnie spytano tych biednych, porażonych prądem ludzi, czy kiedykolwiek chcieliby powtórzyć to doświadczenie, a oni oczywiście powiedzieli, że nie, nigdy, że zdecydowanie było to zbyt nieprzyjemne. A potem spytano, czy w takim razie wolą obejrzeć ostatni film z Martą Żmudą-Trzebiatowską, a oni bez wahania mówili, że no dobra, to dawajcie już ten prąd.

Nie no, OK, to ostatnie zdanie to żart. Niemniej jestem przekonana, że gdyby tylko Amerykanie znali Martę Żmudę-Trzebiatowską, to dokładnie tak by to wyglądało. Profesor Wilson zrobił jednak coś zupełnie innego – zamknął każdą z osób w oddzielnym pokoju o wystroju takim, rzekłabym, nienachalnie ekstrawaganckim: krzesło, ciemne ściany i jakiś stolik. Ochotnikom powiedziano, że muszą pozostać w tym pokoju przez z góry określony czas i robić... nic. Był tylko jeden szkopuł – w pokoju znajdował się jeszcze jeden niewielki przedmiot, umieszczony tuż obok krzesła, na którym siadał badany: ten magiczny przycisk, który raził prądem.

Timothy Wilson chciał bowiem sprawdzić, czy jego ochotnicy osiągną poziom znudzenia tak wysoki, że aż z tego wszystkiego zaczną razić się prądem. Myślicie, że z tych nudów, dla zabawy, faktycznie to robili? Mhm. Średnio dwa razy. Przed upływem siedmiu minut. 2/3 mężczyzn co najmniej raz. No i oczywiście zawsze znajdzie się jeden człowiek, który porazi się 192 razy i trzeba przerwać eksperyment. Wiadomo.

A pamiętacie, jak pisałam, że uczestnicy badania wiedzieli, ile czasu będą musieli się nudzić? Zgadnicie, jak długo? Mhm.

15 minut.

O mój Boże, przecież jak ja to usłyszałam, to natychmiast pomyślałam sobie, że kiedy moi studenci nudzą się na wykładzie, to to jest dokładnie tak, jakbym raziała ich prądem. No właśnie – moi studenci. Musicie wiedzieć, że większość swojego zawodowego życia spędziłam za granicą, bo gdy kiedyś wpisałam w Google: MAM MAGISTRA NIC NIE UMIEM SZUKAM PRACY, to ten poradził mi, żebym wyjechała do Irlandii. Tam pierwszą robotą, do której aplikowałam, była posada rozdawcza balonów w świątyni hamburgerów i cholesterolu i, co ciekawe, opis obowiązków był w miarę adekwatny do nazwy posady, albowiem słuchajcie, ta praca polegała na tym, że się rozdawało dzieciom balony. Oprócz tego, że miałam wszelkie kompetencje do wykonywania kluczowych czynności wymaganych na tym stanowisku, to czułam się szalenie związana emocjonalnie z moim przyszłym pracodawcą, a zwłaszcza z jego cheeseburgerami, więc aplikowałam. Naszkicowałam CV, trzasnęłam list motywacyjny, trzeba było wypełnić bardzo dużo formularzy i testów, więc je wypełniłam, a kilka dni później otrzymałam odpowiedź, że niestety nie otrzymam pracy na stanowisku rozdawcza balonów dzieciom, bo oblałam wymagane testy psychologiczne. No i właśnie dlatego zostałam nauczycielką.

Tym samym dość szybko przyszło mi się przekonać, czy to rzeczywiście jest tak, że większość ludzi staje w płomieniach na widok liczb. Dokładniej rzecz ujmując, przekonałam się wtedy, kiedy na mojej uczelni po raz szósty w tygodniu wybuchł pożar.

Pożar i dramatyczne losy człowieka-krześla

Bo musicie wiedzieć, że jeśli chodzi o hobby i sposoby spędzania czasu wolnego, to nic nie cieszy Irlandczyka bardziej niż możliwość spędzenia kawałka dnia na miłych dla ciała, ducha i ucha czynnościach związanych z fałszywym alarmem przeciwpożarowym. Irlandzkie alarmy przeciwpożarowe są wspaniałe, bo one wybrzmiewają taką melodią, którą przyrównać można do wycia susła przed mutacją – i to takiego susła, co to nie jest zbytnio utalentowany muzycznie; w sensie, że kolędę to jeszcze może zanuci, ale do Backstreet Boys to już go nie chcieli przyjąć.

Najlepiej to było wtedy, gdy nam się w bloku, w którym mieszkaliśmy, takowy alarm zepsuł i wył bez przerwy od trzeciej w nocy do dziesiątej rano, to znaczy tak mi opowiadali sąsiedzi, bo tego dnia o piątej rano to ja

już byłam w robocie, gdzie siedziałam pod biurkiem w pozycji embrionalnej, kiwałam się na boki i płakałam ze wzruszenia z powodu panującej dookoła ciszy. Na pewno zastanawiacie się, co w tamtym momencie robił mój mąż. Otóż odpowiadam: spał. Znaczy, czaicie, ten alarm to tak ija, ija, ija, ale nie tak, że cichutko ija, ija, ija, jakby mama szeptała wam czule do uszka, że wstawaj Mariuszk, kotlety na stole, tylko tak ekstremalnie irytująco IJA, IJA, IJA. A w tym czasie mój mąż spał kamiennym snem i przebudził się tylko po to, żeby mi powiedzieć, że jeśli coś się na serio pali, to żebym wzięła kij od miotły i nadziała na niego kiełbasę, to przynajmniej, kiedy już wygramy tę walkę o życie, będziemy mieli niewielki poczęstunek.

Ja w ogóle uwielbiam samą ideę fałszywych alarmów przeciwpożarowych, bo na przykład u mnie w robocie to wyglądało tak, że za pierwszym razem, kiedy włączył się alarm, to człowiek wybiegał na zewnątrz podekscytowany jak czyżyk przed godami, a następnie obdzwaniał całą rodzinę, że oto ledwo wyrwał się ze szponów śmierci, całe życie stanęło mu przed oczami, mój Boże, on dopiero teraz zrozumiał, że w życiu to trzeba carpe diem, memento mori, bon ton, savoir-vivre i habemus papam. No a kilka alarmów przeciwpożarowych później to na sam ich dźwięk przewracał oczami, jakiegoś maila sobie trzasnął, szedł jeszcze do kuchni, żeby mleko ubić do swojego capuccino, a następnie ruchem powolnym, nienachalnym opuszczał budynek, po drodze tylko dwukrotnie zatrzymywał się obok maszyny z batonikami. No i raz na siku.

Oczywiście inaczej rzecz się miała, gdy alarm przydarzył się w trakcie wykładu, bo wtedy to w mojej sali odbywała się *Jaka to melodia?* połączona z olimpiadą. Moi studenci to już po pierwszej nutce byli na zewnątrz uczelni i biegli tak szybko, ale to tak szybko, że ja to tylko raz w życiu byłam człowiekiem, który tak szybko biegł, i to było wtedy, gdy nad morzem mewa chciała mi ukraść frytkę.

No i na przykład kiedyś na zajęciach opowiadałam im jakąś ekstremalnie ciekawą historię, nie wiem, o czym dokładnie, ale skoro było to superciekawe, to najpewniej chodziło o współczynnik determinacji, bo z niego to jest taki celebryta statystycznego świata, rzecz by można – Beata Kozidrak matematyki. I ja tak moim studentom opowiadałam o tym czymś superciekawym, i wtem alarm! No i jak oni usłyszeli alarm, to trzy sekundy później już nikogo nie było w sali, wszyscy stali na zewnątrz zbici w jedną nieokreśloną masę jak najwspanialsze w życiu bierki, a alarm wył, ija, ija,

ija. No to po chwili ja też stałam wśród tych studentów, w samym swetrze, bo z jakiegoś zaskakującego powodu okazało się, że miłość moich studentów do mojej osoby kończy się tam, gdzie zaczyna się zagrożenie ich życia, co oznacza mniej więcej tyle, że nikt nie chciał wrócić do płonącego budynku po moją kurtkę.

Niemniej po chwili dotarła do nich groza tej sytuacji. Zaczęli szeptać między sobą nerwowo, rozglądać się na boki, aż w końcu wydelegowali jakiegoś ochotnika bohatera, który klepnął mnie w ramię i – w połowie w strachu, a w połowie we łzach – spytał: Janina, czy w sali przypadkiem nie została... – tu wziął przerwę na oddech, łzy przełknął – ...lista obecności?

Lista obecności! Lista obecności!!! Zdenerwowałam się. Człowiek tutaj balansuje na krawędzi życia i śmierci, ledwo wygrał z kostuchą pojedynk na pięści, wywalczył dodatkowe sekundy swojego istnienia, a ich interesuje tylko lista obecności, zero szacunku do kruchej wydmuszki ludzkiego bytu, do cudu życia, który to cud się dokonuje na naszych oczach każdego dnia!!!

A co by było – pytam ich – gdyby w tej sali został jakiś ich kolega, dajmy na to kolega Paraic, bo ten by miał największe prawdopodobieństwo pozostania w płóącym budynku, bo nawet gdyby chcieć go zawołać, to w sumie nie do końca wiadomo jak⁵. No i ten Paraic, opowiadał im, akurat w trakcie moich zajęć postanowiłby się założyć z kolegą obok, że cały się zmieści pod krzesło. No totalnie tak by mogło być – on wszedłby cały pod to krzesło i utknąłby, a kiedy zaczęłyby wyć ten alarm, tak wyć, że ija, ija, to Paraic próbowałby uciekać, rzuciłby się do drzwi jak bażant w krzaki na widok kojota, ale nie zmieściłby się we framudze, bo akurat byłby człowiekiem krzesłem. No i ten budynek tak by płonął i płonął, ale nie możemy wykluczyć, że zaraz zrobi BOOM, a tymczasem kolega Paraic wciąż mógłby siedzieć w sali z tym krzesłem na plecach i nikt o nim nie myśli, nikt! Im tylko ta lista obecności w głowie, a tymczasem ich rówieśnik, bliźniak studenckich doświadczeń, no dobrze, nie bójmy się tego słowa, ich PRZYJACIEL Paraic mógłby właśnie tam płónać, aaaaa, aaaaa! Mój Boże, jak on by płónął, niech ktoś zatrzyma tę spiralę fizycznego cierpienia i rozpadu ducha!

Czy wy w ogóle macie jakieś uczucia, ludzie?! – kończę swoją opowieść wzburzona. Nikogo nie wzrusza dramatyczny los Paraica, który zaraz zginąłby bohatersko tylko dlatego, że chciał udawać krzesło?!

Zamilkli. Popatrzyli na mnie jak gołąb na kromkę suchego chleba, oczy mieli jak spodki, nie miałam żadnych wątpliwości, że w środku płynęły potoki rozgoryczenia i łez spowodowane dramatycznym losem kolegi Paraica. I oni tak patrzą na mnie, zastanawiają się, myślą. W końcu jeden przemówił: Noooooo, to zależy – zaczął ostrożnie – czy kolega Paraic miałby przy sobie listę obecności.

I w ten właśnie sposób dowiedziałam się, że trudno będzie przekonać moich studentów, by nie stawali w płomieniach na widok liczb. Co więcej, okazało się, że są gotowi to zrobić nawet dosłownie, jeśli tylko w grę wchodzi lista obecności. I wiem, co sobie myślicie. Myślicie sobie, że was ten problem ograniczonego afektu do statystyki nie dotyczy, jeśli tylko nie nazywacie się Paraic i nie lubicie w wolnym czasie udawać mebli czy innych elementów wystroju wnętrza. Mam jednak dla was złą wiadomość: ten problem dotyczy wszystkich, bez wyjątku. A nauczył nas tego pewien szalenie emocjonalny kot miłujący wyszukaną muzykę.

Kot, który kochał tureckie disco

Trzy lata temu w serwisie YouTube pojawił się niesamowicie rozczulający filmik, który powodował, że serca wszystkich widzów rozpuszczały się jak smalec na patelni. Film przedstawiał kota oglądającego nagranie, na którym widać było jego niedawno zmarłego właściciela.



Jeśli obejrzyjecie ten film, który wciąż krąży po sieci⁶, to przekonacie się totalnie, że to dzieło sztuki filmowej jest kocią odpowiedzią na *Titanica*. Kot ogląda film, przytula się do telefonu, później przykłada jeszcze swoje

puchate poliki do ekranu... i generalnie jak pokazałam ten film na ostatniej wigilii, to wszyscy odwracali zawstydzeni głowy i mówili, że to nie łązy, to deszcz. Film ten w ciągu dwóch lat zdobył 42 miliony odsłon i ogólnie nie byłoby w tym nic złego, gdyby nie to, że tak naprawdę przedstawia on zupełnie inną historię, to znaczy, owszem, kota. Ale kota, który ogląda żółwia. Słuchając przy tym tureckiego disco.

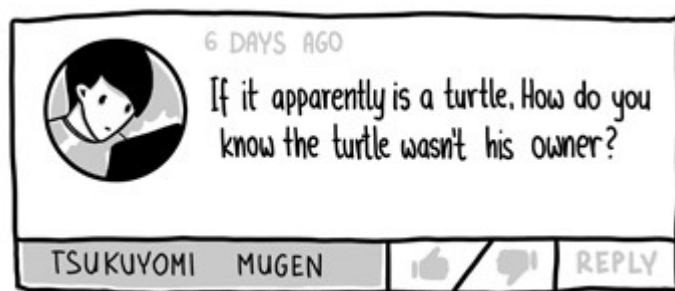
Wprawdzie znalazł się człowiek, który postanowił zabawić się w bojownika prawdy i napisać wszystkim pod owym filmikiem, że ej, ludzie, przecież ten kot ogląda żółwia, a nie swojego właściciela.



Jednak nie wszystkich to przekonało, no bo totalnie wiadomo, że nie, ogląda swojego właściciela.



Szczęśliwie jednak internet nigdy nie zawodzi, więc i tutaj przyszedł z pomocą w postaci genialnego rozwiązania. Oto pojawił się człowiek o odwadze tak wielkiej, że najpewniej byłby w stanie rozsądzić nawet miłośników pomidorowej z ryżem i tych, co jedzą ją z makaronem (?!). Jeśli to żółw... – zastanawiał się nasz bohater, bojownik pokoju i prawdy, użytkownik Tsukuyomi Mugen – jeśli to żółw, to skąd wiemy, czy przypadkiem żółw nie był po prostu właścicielem tego kota?



I o ile w przypadku żółwia opiekującego się kotem pomyłka internautów raczej nie ma większego znaczenia, to dokładnie ten sam mechanizm konsumowania i udostępniania treści dotyczy innych, znacznie poważniejszych tematów. Na przykład takich, o których można przeczytać w nagłówkach niektórych portali:

Wzrost ceny bekonu zwiększa liczbę rozwodów
Życie w biedzie zwiększa rozmiar penisa
USG wykonane w ciąży powoduje leworęczność (ale tylko u chłopców!)
Wszyta pod skórę cieciora leczy raka
Zamordowani holistyczni lekarze odkryli powodujący raka enzym
dotawany do wszystkich szczepionek

Jak widzicie – im dalej, tym straszniej. Bo jeśli zastanowimy się nad tym, z jakich elementów składa się cykl życia badania naukowego, faktu czy dowolnej statystyki, to wygląda to mniej więcej tak: zaczyna się od jakiegoś badania naukowego zrealizowanego (najczęściej) rzetelnie⁷ przez określonych badaczy lub całe instytuty intelektualnej rozkoszy. Jeśli wyniki tego badania są ciekawe (nie tylko dla naukowców) albo kontrowersyjne, to najpewniej temat przedostanie się do mediów, gdzie opanieruje się go w odpowiednią dozę dramaturgii i niedopowiedzeń, a całość posypie szczyptą clickbaitowego nagłówka. I teraz główny problem polega na tym, że niewielu czytelników dotrze do źródła. Większość przeczyta tylko artykuł w porannej gazecie, a niektórzy sam nagłówek, co nie zmienia tego, że wszyscy będą o tym dyskutować. To, w jaki sposób przekazywane są fakty w tym łańcuchu: badania naukowe – media – opinia publiczna, przypomina trochę grę w głuchy telefon, kiedy pierwsza osoba mówi, że lubi brukselkę, a ostatnia słyszy, że mleko od łaciatych kóz powoduje jąkanie.

Słuchajcie, ja zawsze mówiłam, że w życiu poprzeczkę należy sobie zawieszać wysoko, bo wtedy, przechodząc pod nią, nie trzeba się schylać. I właśnie dlatego teraz uczę statystyki i metodologii badań, która – wbrew studenckim ewaluacjom – jest szalenie interesującym przedmiotem. Bo to nie tylko liczby, wzory długie jak spaghetti i zakręcone jak słoik po ogórkach, to przede wszystkim zestaw reguł wnioskowania, na podstawie których możemy ocenić, czy dane badanie naukowe, wnioski logiczny lub sposób myślenia są prawidłowe. Wszystko, co przeczytacie w tej książce, sprawi, że już nigdy nie dacie się omamić spektakularnym nagłówkom w gazecie, na przykład takim, które donoszą, że „w zeszłym miesiącu 2 135 000 Amerykanów zażywało kokainę” (co brzmi zdecydowanie bardziej dramatycznie niż równie prawdziwy tytuł, który oznajmia, że w zeszłym miesiącu 0,7% Amerykanów zażywało ten narkotyk).

A to, by nie dać się oszukać, jest ważne, bo decyzji, czy zaufać konkretnym liczbom, danym, doniesieniom medialnym, podejmujemy codziennie mnóstwo, często nawet tego nie zauważając. Co zresztą najlepiej wyjaśni nam mordercze samochody.

Porozmawiajmy o samochodach z licencją na zabijanie

W ostatnich latach zdarzyły się co najmniej dwa głośne wypadki z udziałem samochodów autonomicznych. Niestety oba śmiertelne. Podejście mediów do tych wypadków było różne. W jednej z gazet mogliśmy przeczytać, że samochody autonomiczne już decydują, kogo zabiją (oryg. *self-driving cars are already deciding who to kill*), co w sumie brzmiało wystarczająco strasznie, by zdecydować się na natychmiastową przeprowadzkę do lasu, a tam granie do końca życia w warcaby z kolegami zrobionymi z patyka i garści kurzu. Drugi nagłówek dotyczący tego samego wydarzenia wyglądał trochę inaczej i donosił, że owszem, autonomiczny uber zabił pieszego, lecz ludzie za kierownicą zabijają dziś szesnaścioro (oryg. *a self-driving uber killed a pedestrian. Human drivers will kill 16 people today*). Oba nagłówki dotyczą tej samej sytuacji, niemniej zostały zaplanowane tak, by wywołać skrajnie różne emocje i opinie. A wiecie, który z nich jest prawidłowy? Żaden.

Bo aby móc tak naprawdę zastanowić się nad tym, na ile zabójcze są samochody autonomiczne i ile tych patyków musimy nazbierać, żeby starczyło nam przyjaciół do końca życia, musielibyśmy wziąć pod uwagę

proporcje między śmiertelnymi wypadkami spowodowanymi przez samochody autonomiczne a tymi, które zostały spowodowane przez ludzi za kierownicą. No i wyliczyć średnią zużywalność przyjaciół z kurzu i patyków. Dopiero takie porównanie dałoby nam odpowiedź na pytanie, które z samochodów są większym zagrożeniem na drodze. A to nie koniec ciekawostek! Teraz czas na nowinkę, która – gdyby była owocem – byłaby najdorodniejszym awokado w supermarkecie nudnych informacji. Ona sprawi, że będziecie brylować na absolutnie każdej prywatce (chyba że wciąż mówicie „prywatka”, to wtedy nie, bo i tak nikt was nigdzie nie zaprasza i wszyscy wiemy, że waszego kolegę Kamila tak naprawdę wymyśliła wam mama – bo w mieście trudno o patyki).

A więc uwaga, ciekawostka! Otóż gdy ostatnio czytałam statystyki motoryzacyjne z dziewiętnastowiecznej Anglii, to dotarły do mnie dwie rzeczy:

1. Zrozumiałam, dlaczego nie mam przyjaciół.

2. Dyskusja, którą toczyliśmy obecnie na temat bezpieczeństwa samochodów autonomicznych, jest identyczna jak ta, która miała miejsce, gdy przesiadaliśmy się z koni do samochodów. Z tym że wtedy o statystykach śmiertelności donoszono poprawnie, a mianowicie raportowano, że 1 na 17 000 przejazdów konno kończy się śmiercią, a w przypadku samochodów ta proporcja wynosi 1:30 000. I co, interesujące, nie? Jestem przekonana, że kolega Kamil byłby pod dużym wrażeniem.

I teraz słuchajcie, ja wszystko rozumiem, ja wiem, co sobie myślicie: że Kamil nie może być pod wrażeniem, bo został zmyślony przez mamę. A także, że przecież takie uproszczone myślenie o świecie, które często spotykamy w serwisach informacyjnych czy dyskusjach internetowych, jest znacznie przyjemniejsze niż poświęcanie czasu na dokładne zapoznawanie się z liczbami i statystykami (to też prawda). Ten błąd poznawczy i wnioskowania nazywamy ILUZJĄ GRUPOWANIA, która oznacza ludzką tendencję do szukania uproszczonych rozwiązań i do łączenia kropek tam, gdzie nie ma żadnego wzoru. Niemniej Hans Roslin, absolutnie cudowny profesor statystyki i popularyzator nauki, powiedział kiedyś coś, co właściwie oddaje cały sens tego rozdziału (nie wliczając w to rzecz jasna wybitnej anegdotki o człowieku-krześle). Powiedział mianowicie: *Let my dataset change your mindset* – niech moje dane zmienią twój sposób myślenia. To liczby i wyniki badań powinny zmieniać myślenie i opinie

społeczeństwa, a nie odwrotnie. Oczywiście liczby, zupełnie tak jak samochody autonomiczne, czasem się mylą. Ale wciąż znacznie rzadziej niż ludzie.

Niech moje dane, statystyki, wyniki badań zmienią wasz sposób myślenia. Jakie to ładne. I jakie... trudne. Zdaję sobie sprawę, że kto by miał czas zastanawiać się nad takimi rzeczami, skoro w internecie jest do obejrzenia tyle kotów z chlebem na głowie i tyłu ludzi, którzy rzucają sobie plasterkami sera w twarz. Rozumiem też, że ludzkość ma inne, znacznie bardziej palące problemy, z którymi od lat nie potrafi sobie poradzić. Na przykład: czy lepiej mieć kaczkę wielkości konia, czy konia wielkości kaczki? Albo: jeśli pływając, spala się 500 kcal na godzinę, to dlaczego wieloryb wciąż jest gruby? No i to jedno kluczowe, absolutnie najważniejsze pytanie: sweter w romby – zbrodnia na modzie czy na geometrii? Niemniej chwilowo zaczniemy od czegoś prostszego. Zastanówmy się, na ile właściwie możemy tej nauce wierzyć, na ile jej ufać. To znaczy – czy na pewno naukowcy są jak żony i nigdy się nie mylą?

Czy naukowcy są jak żony i nigdy się nie mylą?

Wraz z chwilą, w której objęłam stanowisko żony, posiadałam niezwykłą umiejętność, która polega na tym, że wiem wszystko na każdy temat, a mój mąż to nawet uważa, że ja wiem zawsze wszystko najlepiej, choć z jakiegoś powodu twierdzi też, że to nie jest komplement. Dla tych, którzy nie wiedzą, o co chodzi, wyjaśnię. To po prostu działa tak, że wraz z momentem pojawienia się obrączki na palcu nabrałam głębokiego przekonania, że każda z moich opinii jest niezbędna dla istnienia wszechświata, a im mniej wiem na dany temat, tym bardziej kluczowe są moje komentarze. Jeśli zastanawiacie się, gdzie można trenować takie kompetencje, to podpowiadam, że najwięcej profesjonalnych trenerów tej dyscypliny znajdziecie w instytucji zwanej internetem.

Oczywiście czołowym beneficjentem wszelkich złotych rad mojego autorstwa jest mój mąż. Co więcej, Wojtek ma pakiet premium z opcją zapętlenia, to znaczy, że jak on będzie parkował, to ja w trakcie tego parkowania poinformuję go sto razy, że nie ma szans, żeby się zmieścił na tym miejscu, nie zmieści się, no nie ma cudów – nie zmieści się, Wojtek, ale przecież ty się nie zmieścisz, i generalnie to ja tak będę powtarzać do czasu, aż Wojtek się zmieści, wyłączy silnik i wyjdzie z samochodu. I ja wtedy nigdy nie mam pewności, czy jeszcze kiedykolwiek wróci.

No i słuchajcie, ostatnio się przeprowadziliśmy, na przedmieścia się przeprowadziliśmy, do takiego miłego miejsca, gdzie koty siedzą na płotach, dzieci bawią się na ulicy i jeszcze jeździ tam taka grająca ciężarówka sprzedająca lody, którą uważam za wynalazek szatański, no bo to niby lody, czyli pycha, ale żeby do tej rozkoszy dotrzeć na czas, to trzeba galopować jak rozhisteryzowany żrebak, a ja to jednak nie mam aż tyle determinacji, żeby ryzykować bajpasy dla truskawkowego rożka.

W każdym razie nasz proces przeprowadzki wymagał wielu decyzji: co spakować, jak przenieść, gdzie znaleźć, mój Boże, toż to był Disneyland dla żon! I ja się bardzo starałam, naprawdę, no ale to totalnie nie moja wina, że jak na przykład mój mąż wieszał obrazek, to wieszał go krzywo i ja mu mówiłam, że Wojtek, ale to krzywo jest, jak taka zabawka na baterie powtarzałam, że krzywo, krzywo jest, krzywo ten obrazek, a wtedy Wojtek mówił, że uspokój się, kobieto, przecież on jeszcze nie zaczął niczego wieszać, on tylko sięga po młotek, a ja mówiłam, że wszystko spoko, ale

jak już sięga po młotek, to niech jednocześnie zwróci uwagę na to, że ten obrazek to jest krzywo i czy on mógłby go przesunąć bardziej w lewo, w lewo, w lewo, W LEWO! Aha, nie, dobra, w to drugie lewo.

A kiedy już udało nam się zawiesić prosto wszystkie obrazki, rozpakować pudła i podrzeć pozew rozwodowy, to powiedziałam Wojtkowi, że w nagrodę mogę z nim obejrzeć ten film o kosmonaucie, który od dawna bardzo chciał obejrzeć. I obejrzelśmy, i słuchajcie, ten kosmonauta robił jakieś totalnie niestworzone rzeczy: skakał po meteorach, wirował w czarnej dziurze, przenosił się w czasie, a na końcu oczywiście udało mu się uratować cały świat przed zagładą. Jezu – pomyślałam – co za absurdalny film.

Jezu, co za absurdalny film – skomentował mój mąż, jakby czytał mi w myślach – Niby jakim cudem to wszystko mu się udało... tak bez pomocy żony?!

I właśnie tak to jest.

Poświęćmy zatem chwilę na refleksję nad wszystkimi kosmonautami tego świata, którzy musieli sobie radzić bez pomocy (powiedzmy to sobie – kluczowej) wszelkich życiowych klakierów tego świata. No dobrze, zawsze znajdzie się jakiś fanatyk, który stwierdzi, że jeśli wysyłamy kogoś w kosmos, to przecież to przedsięwzięcie ma poparcie w najnowszych odkryciach nauki – inżynierii, robotyce i wielu innych trudnych dyscyplinach, o których nie mamy zielonego pojęcia. Pytanie brzmi: czy naukowcy, zupełnie jak żony, też się (nigdy!) nie mylą? Wszak, jak donosi pewien użytkownik internetu, Ojciec_Mateusz53, którego kompetencje i wiarygodność są dodatkowo wzmocnione przez całkowitą anonimowość: „nauka nie jest nigdy niczym dobrym, bo wprowadza tylko zamieszanie, a de facto jest bardzo ograniczona i ułomna, więc nie warto się nią posługiwać”. No bo – wtóruje mu Marcin B. na LinkedInie: „w czasach, gdy poziom zaufania do wszelkich badań, rankingów, sondaży jest niski, komu należy ufać? Chyba tylko zdrowemu rozsądkowi”. Zresztą – uzupełnia Lubie_maslo1234 – „po co dowody naukowe – wystarczy się po znajomych rozejrzeć, w rodzinie popytać, po mieście pospacerować”. No to rozejrzyjmy się, pospacerujmy po mieście, dokładniej rzecz ujmując: po dziewiętnastowiecznym Wiedniu.

Czym urażono uczucia noworodków?

Ignaz Semmelweis był lekarzem położnikiem, który na początku XIX wieku przybył do Wiednia, by podjąć pracę w głównym stołecznym szpitalu. Nie zdążył nawet jednak porządnie się rozpakować, gdy napotkał problem, i to bardzo poważny – dotyczył on śmiertelności okołoporodowej matek. Okazało się mianowicie, że w latach 1841–1846 w owym szpitalu na każde 10 porodów przypadał jeden zgon matki (co na szczęście obecnie jest statystyką, o której już dawno zapomnieliśmy). W roku 1847 liczba ta się podwoiła. Ignaz Semmelweis zachował się wtedy jak totalny hipis – przed wyciągnięciem jakichkolwiek wniosków postanowił spojrzeć na dostępne dane (?!). Dokładniej rzecz ujmując, to na dwa zestawy danych, bo w owym szpitalu były dwie sale porodowe. Ponieważ Semmelweis znał się na medycynie, a nie na copywritingu, to nazwał owe zestawy w sposób zupełnie przebojowy: sala porodowa 1 i sala porodowa 2.

Dane dotyczące liczby porodów i zgonów w obu salach wyglądały tak:

SALA PORODOWA 1

rok	porody	zgony	%
1841	3036	237	7,8%
1842	3287	518	15,8%
1843	3060	274	9,0%
1844	3157	260	8,2%
1845	3492	241	6,9%
1846	4010	459	11,4%
łącznie	20 042	1989	9,9%

SALA PORODOWA 2

rok	porody	zgony	%
1841	2442	86	3,5%
1842	2659	202	7,6%
1843	2739	164	6,0%
1844	2956	68	2,3%
1845	3241	68	2,0%
1846	3754	105	2,8%
łącznie	17 791	1045	5,9%

Jeśli spojrzymy na najbardziej interesującą nas liczbę, czyli procent zgonów, to okaże się, że w sali porodowej 1 współczynnik zgonów od lat był znacząco wyższy (9,9%) niż w sali porodowej 2 (5,9%). W tej sytuacji Semmelweis zadał sobie jedyne słuszne pytanie: co te dwie sale różni?

Warunki sanitarne? Nie, te na obu oddziałach były takie same. Może stan pacjentek – może do pierwszej sali porodowej kierowano znacznie cięższe przypadki? To również nie była prawda, pacjentki przydzielano bowiem całkowicie losowo. A może problem tkwił w personelu pracującym w obu salach? Bingo! Rzeczywiście było tak, że w jednej sali pracowali tylko lekarze, a w drugiej – tylko położne. Ale co ciekawe, w sali, gdzie

pracowali tylko lekarze, śmiertelność była wyższa. No to myślmY dalej: kto był lekarzem w XIX wieku? Wyłącznie mężczyźni. I to właśnie doprowadziło pierwszych badaczy do następujących wniosków, które skrupulatnie zanotowali w dzienniczku badawczym: „Noworodki były urażone koniecznością przyjsia na świat w obecności mężczyzn” (oryg. *Newborns were offended by having been delivered in the presence of men*). Co podobno zwiększało ryzyko wszelkich powikłań u matek, wiadomo, z takim wkurzonym noworodkiem to szalenie trudno współpracować.

Z jakiegoś dziwnego powodu Ignaza Semmelweisa nie przekonało jednak wyjaśnienie o urażonych noworodkach i kopał dalej. Nawet całkiem dosłownie, bo w swoich poszukiwaniach dotarł aż do piwnicy, gdzie znajdowało się prosektorium. Prosektorium, w którym zjawiali się wszyscy lekarze przed dyżurem, by potrenować sekcje zwłok, a zaraz potem udawali się na blok porodowy. Oczywiście nie myli po drodze rąk, bo w tamtych czasach jakoś nikt nie zawracał sobie głowy bakteriami⁸.

Gdy Ignaz Semmelweis wykminił, że winowajcy zwiększonej śmiertelności rzeczywiście są niewielcy, chociaż znacznie mniejsi niż podejrzewane wcześniej wkurzone noworodki (no bo są bakteriami), to śmiertelność na obu oddziałach zmniejszyła się do 1%. Ten przykład fajnie pokazuje, że w logice i trudnej sztuce wyciągania wniosków jest coś znacznie ważniejszego od tych naszych tabelk i komórek Excela – mianowicie nasze szare komórki. A najlepszą przyjaciółką tego zestawu komórek jest kolejna słodka rodzinka w serniku nauki⁹ – metoda naukowa.

Skąd wiemy, czy wielkość penisa wpływa na rozwój gospodarczy kraju?

Metoda naukowa to sposób dotarcia do prawdy i sformułowania prawdziwego opisu zjawiska. Innymi słowy – zespół zasad i reguł, które pozwalają nam jak najrzetelniej odpowiedzieć na pytanie badawcze, które nas trapi. Na przykład takie, które swego czasu dręczyło Fina Tatu Westlinga (2011), to jest czy istnieje związek pomiędzy średnią wielkością penisa w danym kraju a jego rozwojem gospodarczym? Żeby znaleźć wiarygodną odpowiedź na to pytanie, musielibyśmy przejść przez poszczególne składniki metody naukowej (zwane również etapami pracy naukowej).

1. Przegląd literatury

To zupełnie nie jest tak, że jak taki Tatu Westling sobie wymyśli, że zrewolucjonizuje światową gospodarkę za pomocą męskiego przyrodzenia, to wystarczy, że spakuje do walizki linijkę i dużo uroku osobistego, a następnie ruszy na ulice prosić napotkanych mężczyzn, żeby zmierzili sobie penisa i jednak nie dzwoniли po policję. To badanie, jak każde inne, powinno zacząć się gdzie indziej – od przeglądu już istniejącej literatury naukowej na interesujący nas temat. Dzięki temu dowiemy się, jaki jest stan wiedzy w danej dziedzinie, jakimi teoriami naukowymi wcześniej próbowano to zagadnienie wyjaśnić, a także czy istnieją w tej kwestii jakieś punkty sporne i dyskusje między dyscyplinami i naukowcami.

Bo to oczywiste, że naukowcy czasem się kłócą. To znaczy – najpierw jeszcze udają, że nie, na początku to jest niby miło, niby dzień dobry i co słyhać u żony, a potem nagle następuje jakaś prowokacja, na przykład ktoś pyta kogoś, czy może otworzyć okno, a to działa jak płachta na byka, i stąd to już tylko rzut beretem do chaosu i festiwalu przemocy, chwilę później już wszyscy są na komisariacie i ten jeden naukowiec tłumaczy panu policjantowi, że owszem, uderzył kolegę łopata¹⁰, ale tylko dlatego, że ten powiedział, iż dyscyplina, którą tamten uprawia, nie jest prawdziwą nauką, a generalnie to i tak cud, że trafił, bo ten pierwszy ochraniał się krzesłem topornym jak jego własna metodologia. No i był jeszcze z nimi jeden kolega z wydziału, ale ten to akurat jest na oddziale ratunkowym, bo zjadł doktorat tego pierwszego, by udowodnić mu, że do niczego innego się ta praca nie nadaje.

Ale jak już naczytamy się literatury i wypuszczą nas z SOR-u, po konsultacji z gastrologiem i być może psychiatrą, to możemy przejść do drugiego etapu każdego badania naukowego.

2. Sformułowanie pytań badawczych

Pytanie badawcze to konkretne pytanie, na które badacz próbuje odpowiedzieć w toku swojej pracy naukowej. W przypadku Tatu Westlinga mogło ono brzmieć na przykład tak: jaki jest związek między średnią wielkością penisa w danym kraju a jego rozwojem gospodarczym? Matka Tatu Westlinga mogłaby z kolei sformułować je w następujący sposób: co zrobiłam źle?!

3. Skonstruowanie hipotez badawczych

Hipoteza to naukowa propozycja odpowiedzi na pytanie badawcze. Takie szczegółowe twierdzenie, które następnie będzie testowane w toku badania. Hipotezy badawcze muszą więc być weryfikowalne, co oznacza, że musi istnieć jakiś sposób na sprawdzenie, czy są prawdziwe, czy fałszywe. Technicznie rzecz ujmując, nawet stwierdzenie, że Księżyc jest zrobiony z sera, jest hipotezą badawczą. Przynajmniej od czasu, kiedy możemy na ów Księżyc polecieć, pobrać odpowiednie próbki gleby, a następnie usiłować zrobić sobie z nimi grzanki.

4. Pobranie próby badawczej

To oznacza, że musimy zdecydować, w jaki sposób dobierzemy osoby, które wezmą udział w naszym badaniu. Więcej o doborze próby piszę w rozdziale *Dlaczego koń i szop pracz nie dostały doktoratu?*, na razie warto wiedzieć, że to nie do końca jest tak, że jak sobie przygotujemy jakąś ankietę, a później rozdamy ją mamie, sąsiadce i jej trzem szynszylom, to już będzie się to liczyło jako wiarygodne badanie naukowe. Zwłaszcza że szynszyle nie mają kciuków, więc mogą być dość sfrustrowane podczas pisania, co z kolei sprawia, że mogłyby wypełnić ankietę na odwal się.

5. Zebranie danych

Tutaj możliwości jest mnóstwo! Czy Tatu Westling postanowił sam dokonać pomiarów? Czy skorzystał z danych dostępnych w narodowych statystykach medycznych (jeśli takowe istnieją)? Czy rozdał wszystkim ankietę i pozwolił samemu wpisać wynik? Metoda, którą wybierze badacz, powinna być ściśle związana z jego pytaniem badawczym, a także możliwościami praktycznymi, finansowymi i czasowymi.

6. Analiza danych

Tylko uwaga – z tą analizą danych to jest trochę szkopał, bo gdy piszę o analizie danych, to nie mam na myśli tak popularnej w naszym kraju metody analizy na czuja, ewentualnie równie popularnej metody na szwagra („mój szwagier miał intuicję”).

Słuchajcie, Intuicja to dobre imię dla jamnika, ale niekoniecznie dobra metoda wyciągania wniosków. Tym bardziej, że musicie wiedzieć, iż metoda naukowa ma również rodzeństwo, dokładniej rzecz ujmując, brata.

I to takiego, którego niechętnie zaprasza się na wszelkiej maści schadzki rodzinne.

Zły bliźniak metody naukowej

Zły brat bliźniak metody naukowej jest jak człowiek, który wchodzi do świeżo posprzątanego salonu prosto po kąpieli w pobliskiej sadzawce, nie zdjawszy nawet kaloszy, a na urodzinach ojca podmienia wszystkie bezy na takie wegańskie, robione z wody po cieciorce. Taki, który paraliżuje wszystkie uroczystości rodzinne i zmusza matkę do chowania całej biżuterii i kryształowej cukiernicy, bo wiadomo, że wyniesie pod kapotą, sprzeda, a kasę przepuści na oranżadę i balony o ordynarnych kształtach. Zły bliźniak metody naukowej to jest, proszę państwa, DOWÓD ANEGDOTYCZNY.

To jest tak, że gdyby w końcu jakaś stacja zechciała kupić ode mnie serial o ludziach-statystykach, to byłby to niewątpliwie najstraszniejszy program w telewizji (oczywiście przerywany tymi przezabawnymi momentami, kiedy to główny bohater opowiadałby brawurowe żarty o tym, że przychodzi rozkład t-Studenta do lekarza, a lekarz mówi: „Ale proszę mi się tutaj rozłożyć normalnie”, hi, hi). Byłby to serial kryminalny i pierwszy odcinek wyglądałby tak, że siedzi sobie pan policjant w chałupie, ogląda Ojca Mateusza, żeby się czegoś nauczyć o rozwiązywaniu poważnych zagadek kryminalnych, a tu nagle dryń, dryń, dzwoni telefon, więc on odbiera i głos w słuchawce mówi, żeby przyjeżdżał jak najszybciej, bo było morderstwo, więc on ijo, ijo, jedzie na miejsce zbrodni, patrzy, rzeczywiście – zwłoki. Są zwłoki i jest też lekarz patolog, więc ten policjant go pyta ojojjoj, co tu się stało, a patolog mówi, że statystyk umarł, a policjant pyta, czemu umarł, a patolog mówi, że wie pan, z żalu umarł, bo podczas internetowej dyskusji o faktach naukowych ktoś napisał: „To musi być prawda, u mnie się sprawdziło”.

0101. A CZEMUŻ TEN STATYSTYK UMARŁ?



DOWÓD ANEGDOTYCZNY TRAFIŁ GO W SERCE

A wiecie, gdzie najczęściej można znaleźć takie zdanie? Na forach dla matek i kobiet w ciąży, w dyskusji na temat przewidywania płci dziecka. I tam zawsze znajdzie się jakiś post Heleny z Elbląga, która mówi, że koleżanka listonosza kuzynki jej matki mówiła, że jeśli całą ciążę będzie chodzić tyłem, to będzie mieć syna, czy to prawda? I tam pod tym postem zawsze jest milion komentarzy, że „hej Helena, to prawda, ja całą ciążę chodziłam tyłem i teraz mam chłopca, u mnie się sprawdziło”, „siema Helena, ja w ciąży to tyłem nawet na Giewont weszłam, a teraz mam dwóch synów, u mnie się sprawdziło”, „cześć, Helena, ja chodziłam przodem, ale w ciążę to zaszłam tyłem i też będę mieć syna”.

I oczywiście zawsze znajdzie się jakiś psychopata, morderca wszelkich uciech i radości życia, który stwierdzi, że prawdopodobieństwo posiadania syna lub córki jest stałe dla wszystkich ciąż, nieważne, czy kobieta będzie chodzić przodem, tyłem, czy bokiem. Niestety wszelkie dowody anegdotyczne, dowody w postaci „historyjki o kuzynce szwagra”, są dla nas szalenie atrakcyjne, a to dlatego, że dają uproszczone i szybkie odpowiedzi na większość trapiących nas pytań.

Wszystkie wasze serniki były puchate niczym mały króliczek angora, a ten jeden miał konsystencję cegłówki wymieszanej ze żwirem? Przecież mama mówiła, by sernik mieszać tylko w jedną stronę, bo inaczej się obrazi! Nigdy nie udawało wam się ugotować idealnego jajka na miękko, ale od czasu, gdy robicie to w skarpetkach w kury, jeszcze nie ponieśliście porażki? Od dawna nie jadacie drożdżówek z serem, bo kiedyś jedna z nich zabiła sąsiadkę z parteru? To znaczy owa sąsiadka kupiła drożdżówkę, a potem to zabił ją nadjeżdżający samochód. Ale niesmak pozostał.

Gdy przydarza nam się coś nieoczekiwanego, lubimy szukać prostych wyjaśnień takiego stanu rzeczy. Niemniej to, że coś przydarzyło się nam, a nawet Helenie z Elbląga, nie wystarcza, by uznać to za regułę czy fakt. To, że jedno zdarzenie następuje zaraz po drugim, wcale nie oznacza, że zaszedł tutaj związek przyczynowo-skutkowy. No i wreszcie – czy skarpetki w kury to na pewno JEDYNY czynnik, który uległ zmianie podczas czynności gotowania jajka? A może zmieniliśmy czas gotowania, garnek albo jajka są innej wielkości? Dopóki nie kontrolujemy wszystkich zmiennych, które mogą wpłynąć na wynik naszego eksperymentu – nieważne, czy chodzi o puchatość sernika, czy stopień ugotowania jajka – dopóty musimy zachowywać niezwykłą ostrożność w wysnuwaniu jakichkolwiek wniosków. Zresztą nauka również ma pewien rodzaj poduszki bezpieczeństwa, która chroni ją przed roztrzaskaniem się o Gubałówkę nierzetelnych badań i inne szczyty nieuczciwych eksperymentów.

Dzień dobry, metodologiczna policja, dowodzik poproszę

Jeśli pisaliście kiedyś jakąkolwiek pracę dyplomową – czy to licencjat, czy pracę magisterską, czy też krótkie sprawozdanie z badania, które miało udowodnić waszym rodzicom, że mycie zębów zwiększa ryzyko straszliwych chorób somatycznych, zespołu stresu pourazowego lub nawet chęci studiowania pedagogiki¹¹ – to wiecie, że musieliście bardzo skrupulatnie opisać proces przeprowadzenia waszego badania. Zawrzeć w rozdziale metodologicznym wszystko: kiedy i gdzie zostało przeprowadzone badanie, jak liczna była próba, jakich metod badawczych użyto, w jaki sposób zoperacjonalizowano wszelkie koncepty, jakie definicje przyjęto... oraz inne detale, które mogły mieć wpływ na wynik badania. Wszystko należało opisać ze szczegółami, do tego stopnia, że jeśli użyliście w swoim badaniu ankiety, to kwestionariusz również powinien zostać dołączony do sprawozdania.

To wszystko robimy z trzech powodów: po pierwsze, dzięki temu zwiększamy objętość naszej pracy magisterskiej w sposób inny niż za pomocą fontu o rozmiarze 78 i możemy przestać tłumaczyć promotorowi, że zaprawdę nie liczy się ilość, lecz jakość, i być może ta praca ma tylko trzy strony, ale, proszę pana, jakże rzetelne!

Po drugie, taki szczegółowy opis jest nam potrzebny na wypadek, gdyby miała zjawić się u nas metodologiczna policja i poprosić o dowód na uzyskane wyniki i wyciągnięte na ich podstawie wnioski. Szczegółowy rozdział metodologiczny jest niczym rodowód w przypadku najbardziej rasowych kotów o płaskich twarzach – pozwala niezależnym sędziom ocenić, czy badanie zostało przeprowadzone prawidłowo i czy można mu ufać. W świecie naukowym takich sędziów używamy najczęściej w procesie recenzowania naukowego (*peer review*), które polega na tym, że z tekstem zapoznaje się, po czym ocenia go kilku specjalistów z danej dziedziny. Nie jest to proces idealny – i to nie tylko z powodu tego, że nie występują w nim koty¹² – niemniej w pewien sposób pozwala nam trzymać rękę na miłym pulsie rzetelnej nauki.

Chociaż z tymi kotami to nie jest do końca prawda. Wszak największą zaletą pisania artykułów naukowych jest to, że pozwalają docenić naszych współpracowników i najważniejsze osobistości naszego naukowego życia. Na przykład w 1975 Jack Hetherington napisał artykuł, w którym używał liczby mnogiej („my przeprowadziliśmy”, „my policzyliśmy”, „my stwierdziliśmy”), mimo że był jego jedynym autorem, o czym przypomniał sobie dopiero pod koniec pisania. Mógł po prostu zmienić liczbę mnogą na pojedynczą w całym dokumencie, no ale bez przesady. Wiecie, ile to roboty? Dlatego dodał drugiego autora. Do dziś odwołujemy się więc do wspólnego artykułu Jacka Hetheringtona i F.D.C. Willarda (to naprawdę bardzo ważny tekst dotyczący zachowań atomu). Jest to jeszcze bardziej imponujące, jeżeli weźmie się pod uwagę, że F.D.C. Willard miał siedem lat. I aha, był kotem. Ale jeśli myślicie, że to jedyny fizyk wśród zwierząt puchatych, to jesteście w błędzie. W 2001 roku A.K. Geim (późniejszy laureat Nagrody Nobla za współodkrycie grafenu) opublikował artykuł wraz ze swoim kolegą podpisanym H.A.M.S. ter Tisha. To jest – ze swoim kolegą chomikiem. Sami więc widzicie, ile to daje możliwości. Jeśli pokończyłyvezenty urodzinowe dla waszych zwierzaków (bo ileż można robić tort z kiełbasy przekładany kremem z salcesonu), to śmiało, niech piszą z wami artykuły – może ucieszą się z Nagrody Nobla¹³ ?



No i trzecia rzecz – taka szczegółowa instrukcja dotycząca tego, jak zostało przeprowadzone badanie, umożliwia jego powtórzenie przez innych naukowców, czyli REPLIKACJĘ. Ta zaś pozwala innym naukowcom ponownie przeprowadzić dane badanie i sprawdzić, czy jego wyniki są powtarzalne. Jedna publikacja nauki nie czyni, niemniej replikacja danego badania zwiększa nasze zaufanie do wniosków, które z niego płyną. A czasem pozwala również namierzyć nieuczciwych naukowców, których w historii nauki nie brakuje. Na przykład takich, którzy nadużyli zaufania nie tylko ludzi, lecz także myszy.

Oszukani ludzie, nadszarpnięte zaufanie myszy

We wczesnych latach 70. w prestiżowym Instytucie Sloan-Kettering pracowano nad przeszczepem skóry z czarnej myszy na białą. Choć sam koncept może wydawać się dość dziwny, to tak naprawdę owe próby miały pomóc w rozwiązaniu problemów z odrzutami przeszczepów u ludzi. Przez długi czas eksperymenty kończyły się fiaskiem z powodu różnic genetycznych pomiędzy różnofutrzastymi myszami. Ale do czasu! W końcu pewien naukowiec z rzezonego instytutu, William Summerlin, ogłosił światu sukces transplantacji, pokazując na dowód całe pudełko cętkowanych myszy. Problem był tylko jeden – pomimo wielu replikacji nie udało się nigdy powtórzyć tego wyniku. Oczywiście nie możemy mieć stuprocentowej pewności, istnieje jednak dość duże prawdopodobieństwo, że wszystkie te niepowodzenia wynikały z tego, że czarno-białe myszy, będące dowodem na sukces transplantacji, zostały przez owego naukowca pocętkowane za pomocą... markera. Świat nauki nie był pod wrażeniem. Myszy zresztą też nie. Gdyby istniało jakieś Stowarzyszenie Najgorszych

Przestępców Wszechświata, to William Summerlin prawdopodobnie dostałby od niego honorowy (aczkolwiek niezbyt prestiżowy) tytuł i statuetkę markera wyrzeźbioną z masła.

Takie niechlubne przypadki rzeczywiście zdarzają się w nauce, choć liczne replikacje i proces *peer review* pozwalają je dość szybko wykrywać. Na przykład wtedy, gdy artykuł doktora Thorakkaiego Shamima dotyczący walki z plagiatami musiał zostać usunięty z czasopisma „The Indian Journal of Dermatologists”, ponieważ... okazał się plagiatem. Czasem jednak nawet świat nauki traci czujność, co może mieć znacznie poważniejsze konsekwencje niż to, że naukowiec usłyszy od mamy, że ją rozczarował, a koledzy będą się z niego śmiali. Najśłynniejszy i jednocześnie najtragiczniejszy przykład takiego oszustwa w nauce dotyczy tematu, o którym – jeśli tylko nie żyliście przez ostatnie 50 lat pod kamieniem – na pewno słyszeliście: związku między szczepieniami a autyzmem.

Naukowiec, za którego wszyscy się wstydzimy

W 1998 roku Andrew Wakefield wraz z grupą naukowców opublikował serię artykułów, w których zasugerował związek między trójskładnikową szczepionką przeciwko odrze, śwince i różyczce a autyzmem i zapaleniem jelit. Pierwsze doniesienie brzmiało rewelacyjnie, choć było to tylko wstępne doniesienie z badań, które w ostateczności okazało się fałszerstwem. Główny artykuł z tego cyklu nie spełniał żadnych norm rzetelnego badania naukowego – badanie przeprowadzono jedynie na 12 dzieciach, nie kontrolowano pozostałych zmiennych, które mogły mieć znaczenie dla rozwoju zaburzeń, a co najważniejsze – „wina” szczepionek została udowodniona nie za pomocą metody eksperymentalnej czy modelowania statystycznego, ale... opinii rodziców. To znaczy, wiecie, przychodził taki człowiek do rodzica i pytał: czy myślisz, że przyczyną złego stanu zdrowia twojego dziecka była szczepionka MMR? A ośmioro na 12 rodziców odpowiadało: TAK. Ej, czytaliście kiedyś podręcznik do zadawania sugerujących pytań ankietowych? Teraz już nie musicie, bo ma tylko jedną stronę, na której przytoczono ten właśnie przykład.

Dalej było tylko gorzej – okazało się, że Wakefield skłamał w sprawie otrzymania zgody komisji etyki na przeprowadzenie owego badania (na dzieciach!), a także w kwestii doboru próby – okazał się on celowy, czyli

Wakefield zakwalifikował do badania tylko te dzieci, które potwierdzały postawioną przez autora tezę. A to jeszcze nie wszystko – następnie okazało się, że tylko jedno dziecko z próby rzeczywiście cierpiało na zaburzenia autystyczne (Wakefield pisał o dziewięciorgu dzieciach). Nieprawdą okazało się też to, że żadne z dzieci nie przejawiało zaburzeń przed szczepieniem (pięcioro z nich cierpiało na różnego rodzaju zaburzenia rozwojowe).

A, no i jeszcze jedno: zaburzenia, które – według artykułu – wystąpiły kilka dni po szczepieniu, tak naprawdę ujawniły się u dzieci wiele miesięcy po otrzymaniu szczepionki. Wszystko zaś ułożyło się w naprawdę sensowną całość, gdy okazało się, że Wakefield otrzymał za publikację swojego artykułu 435 000 funtów od prawników, którzy działali w imieniu rodziców domagających się odszkodowań za rzekome szkody wywołane przez szczepionki. Niemniej z pewnym smutkiem muszę odnotować, że to wszystko wiemy nie dzięki wybitnemu zaangażowaniu naukowego świata, ale dzięki dziennikarskiemu śledztwu, które przeprowadził Brian Deer z „The Sunday Times”.

W przypadku artykułu Wakefielda środowisko naukowe początkowo zawiodło i chociaż później ochoczo rzuciło się do naprawiania szkód dzięki finansowaniu kolejnych badań nad negatywnymi skutkami szczepień oraz publikowaniu prostujących artykułów – było już trochę za późno. W Wielkiej Brytanii, a potem na całym świecie zaczął spadać odsetek szczepionych dzieci. Światowa Organizacja Zdrowia donosi, że żeby przerwać transmisję wirusa odry w społeczeństwie i zapewnić ochronę najbardziej zagrożonym grupom (na przykład niemowlętom lub osobom, które z przyczyn zdrowotnych nie mogą zostać zaszczepione), konieczne jest utrzymanie wyszczepialności populacji na poziomie 95%, bo tylko ten poziom gwarantuje odporność zbiorową. Z danych z roku 2017 wynika, że w krajach Unii Europejskiej i Europejskiego Obszaru Gospodarczego tylko cztery państwa spełniają to kryterium (OECD 2019). Nie ma wśród nich Polski.

Sytuacji nie zmieniają nawet kolejne publikacje badań, które potwierdzają bezpieczeństwo szczepionek (m.in. Hviid et al. 2019, Farrington i Taylor 2001, Mrozek-Budzyn et al. 2010)¹⁴. Z pewną dozą przyjemności piszę o tym, że wśród nich jest także jedno trwające sześć lat badanie, które zostało przeprowadzone na zlecenie amerykańskiej

organizacji antyszczepionkowej SafeMinds. Wykazało ono... absolutny brak związku między szczepieniami a autyzmem (Gadad et al. 2015). Ups.

Jeśli oglądaliście kiedyś jakikolwiek film katastroficzny, to wiecie, że one wszystkie przebiegają tak samo – gdzieś tam na początku pojawia się jakiś naukowiec i ostrzega, że zaraz nadejdzie tornado, atak dinozaurów lub atak tornada pełnego dinozaurów, ale absolutnie wszyscy go ignorują. No, o ile tornada T-rexów zdarzają się stosunkowo rzadko, o tyle ignorowanie dobiegającego ze świata nauki głosu rozsądku jest powszechną praktyką.

Dlatego też rosnące statystyki odmowy szczepień mnie martwią, ale nie dziwią – tam, gdzie w grę wchodzi emocje, pierwszym zakładnikiem bardzo często stają się fakty. Wiele doniesień naukowych nie wymaga zaawansowanej wiedzy matematycznej czy specjalistycznej. Zaryzykuję nawet stwierdzenie, że w erze dostępu do bibliotek, wykładów otwartych i internetu edukacja nie jest droga. Ale gdy jej brakuje, to mamy do czynienia z sytuacją taką jak obecnie: w XXI wieku 29% zgonów dzieci do lat pięciu jest wywołanych przez choroby, na które znamy szczepionki (WHO 2014). I to nie tylko dlatego, że nie wszystkie kraje mają do nich dostęp – również dlatego, że wiele lat temu ktoś posłał w świat fałszywą informację, że szczepionki powodują autyzm, a potem sprawa wymknęła się spod kontroli. Tak, edukacja nie jest droga, ale już ignorancja może nas bardzo wiele kosztować.

Na szczęście skoro czytacie tę książkę, to znaczy, że już możecie przymierzać czapkę prymusa, bo miło wyróżniacie się z tłumu ludzi, którzy gorąco wierzą, że wysoka korelacja między spożyciem sera a liczbą zgonów wskutek zaplątania się w pościeli oznacza, że goudę warto jeść z daleka od wszelkich okolicznych kołder¹⁵. Ale zacznijmy od początku, od pierwszej rzeczy, którą należy sprawdzić, gdy weryfikujemy jakiegokolwiek doniesienie naukowe. Pomoże nam w tym kilka puchatych stworzeń. Znacie ten typ: starały się z całych sił, a doktoratu i tak nie zrobiły¹⁶. Nie, nie mówię o sobie, tylko o dwóch koniach i szopie praczu.

Dlaczego koń i szop pracz nie dostały doktoratu?

W 2017 roku 65 000 ludzi obejrzało, jak wyplatałam koszyk z kabanosów w trakcie relacji na Facebooku, co jest jakby ostateczną odpowiedzią na pytanie mojej mamy: „Janina, a co ty będziesz robić na tym doktoracie?”. Nie jestem jednak szalona, nie mam żadnych wątpliwości, że nie wszyscy naukowcy na świecie mogliby zostać beneficjentami splendoru o takim kalibrze. Na przykład: szopy pracze. Po pierwsze dlatego, że rzadko kiedy mają konto na Facebooku. Po drugie dlatego, że kabanosy mogłyby nie dotrzeć do końca nagrania. I po trzecie, bo szopy nie mają ukończonych studiów doktoranckich. Choć przyznać trzeba, że w 1962 roku jeden z nich był blisko otrzymania tego wyróżnienia.

Małżonkowie J.M. i Helen Warren przeprowadzili eksperyment z dziedziny psychologii, w którym weryfikowali teorię zdolności poszczególnych gatunków do uczenia się odwróconego. Co ciekawe, już wcześniej ustalono, że tę umiejętność posiadają szczury (Dufort, Guttman i Kimble 1954) i koty (Cronholm, Warren i Hara 1960), a ryby niestety jej nie mają (Warren i Warren 1962). Co oznacza, że jeśli znajdziesz jakiegoś dorodnego karasia, dasz mu piątaka i powiesz, żeby nie wydał na głupoty, to on wyda wszystko na głupoty, a potem jeszcze poprosi o więcej i wicie, co zrobi z kolejnym piątakiem? NO BYNAJMNIEJ NIE WPLĄCI GO NA KONTO OSZCZĘDNOŚCIOWE!

Państwo Warrenowie nie poprzestali na surowej ocenie rozrzutnego karasia i postanowili zbadać zdolność uczenia się wśród kolejnych gatunków zwierząt. W tym celu przygotowali krótkie zadania do wykonania, a następnie zaprosili na badanie dwa konie, które nazywały się odpowiednio Pancho i Rag Mop, i jednego szopa pracza. Niestety imienia szopa pracza nie podano w oryginalnym artykule, niemniej zawarto tam adnotację, że bardzo lubił jeść świńskie nerki. Serio. I niech mi ktoś jeszcze powie, że nauka nie jest ciekawa!



Cała rodzina szopa była z niego bardzo dumna, jednak udział w rzeczonym badaniu, a nawet publikacja jego wyników nie sprawiły, że udało mu się wspiąć po szczeblach kariery naukowej. Nie zdobył międzynarodowej sławy ani nawet głupiej szansy na Nobla. A jeśli miałabym kiedyś wyjaśnić rodzinie rzeczonoego pracza albo przyjaciółom dwóch koni, dlaczego to badanie nie spotkało się ze znacznie większym uznaniem środowiska naukowego, to musiałabym powiedzieć, że sama metoda była w porządku, wybór puchatych badanych też dałoby się uzasadnić. Niepokój budzi jednak ich liczba.

Bo to jest trochę tak, że gdy przeprowadzamy badanie, na jakikolwiek temat, to musimy wybrać **POPULACJĘ BADANĄ**, czyli grupę osób (albo zwierząt, krajów, instytucji itd.), która w tym naszym naukowym rendez-vous weźmie udział. Wbrew powszechnej opinii populacja badania nie oznacza „wszystkich ludzi na świecie” lub też – w przypadku zwierząt – „wszystkich szopów na świecie i dwóch koni”. To od badacza (i od pytania badawczego) zależy, jak sobie tę populację zdefiniuje. Zastanówmy się nad tym na przykładzie ryb, no bo w sumie dlaczego nie?

Kenneth H. Mann (1965), biolog morski, postanowił zbadać związek między tym, co ryby jedzą, a tym, jak rosną wszcz. Czyli trochę taki eksperyment, jaki przeprowadzają na nas babcie przy okazji każdego niedzielnego obiadu. W tym celu mógł oczywiście ważyć wszystkie ryby na świecie, ale to by było skomplikowane – po pierwsze dlatego, że ryb jest całkiem sporo, więc nikt nie ma czasu ani pieniędzy na to, żeby je wszystkie zważyć, a po drugie, ponieważ takie ryby to się bardzo trudno liczy. No bo człowiek zaczyna: jeden karaś, drugi karaś, sandacz i dwa

karpie, a za chwilę to już nie wie, czy ten łosoś to jakiś nowy, czy też znowu daje pod prąd.

Na szczęście Mannowi z pomocą przyszła metodologia, bo to od niego zależało, jak sobie ową populację ryb zdefiniował – mógł uznać, że są to wszystkie ryby świata albo tylko ryby rzeczne, albo być bardziej szczegółowym i skupić się tylko na jednym określonym gatunku, albo – jak stało się w tym wypadku – na rybach zamieszkających w konkretnej rzece, tj. Tamizie. Z punktu widzenia metodologii wszystkie te opcje są tak samo prawidłowe. Konsekwencją są jednak ograniczenia w możliwości wysnuwania wniosków, to znaczy, że nie możemy wysnuwać wniosków o szczupakach na podstawie badania karasi. Dobrze jednak, żeby za takim wyborem definicji danej populacji stało jakieś uzasadnienie (teoretyczne lub praktyczne), a nie że wybrałem karasie, bo nie wiem, jak się pisze szczupak.

Dla nas, naukowych świrów, oznacza to, że warto wyrobić w sobie nawyk sprawdzania, na kim zostało przeprowadzone każde badanie, które wpadło w nasze ręce – gazety lubią pewne kwestie upraszczać, piszą na przykład, że „przebadano dzieci przedszkolne”, i wszyscy są zadowoleni, a redaktor naczelny klaszcze, ale dla naszego rozumienia i interpretacji wyników istotne jest, czy były to dzieci przedszkolne ze Szwecji, z Warszawy¹⁷, czy może z grupy Tygrysków z jednego konkretnego przedszkola.

No i jest jeszcze jeden problem – spróbujmy sobie wyobrazić, że chcemy przebadać wszystkie warszawskie przedszkolaki albo wszystkie karasie z polskich rzek. POWODZENIA. Przecież to zajmie wieki, zanim człowiek dotrze z ankietą do każdego polskiego karasia, a z przedszkolakami będzie jeszcze gorzej, bo trochę minie, nim nauczą się czytać¹⁸. Najczęściej więc brakuje nam zasobów czasowych i finansowych, by przebadać absolutnie każdego reprezentanta danej populacji – wyjątkiem jest cenzus, czyli powszechny spis ludności, podczas którego ankieterzy próbują dotrzeć z ankietą do absolutnie każdego obywatela i rezydenta danego kraju. Warunki przeprowadzania takiego cenzusu są różne w różnych krajach – w większości przeprowadza się go co 10 lat, choć istnieją kraje, gdzie robi się to co pięć lat (na przykład Australia, Irlandia, Filipiny czy Japonia).

Z ciekawszych rzeczy – w 2014 roku spektakularnie nie powiódł się spis ludności w Indiach, w całym kraju zabrakło bowiem... ołówków do wypełniania kwestionariusza. Stało się tak dlatego, że badacze nie

doszacowali mniej więcej 15% populacji (co – umówmy się – jeszcze bardziej uzasadniało konieczność przeprowadzenia tego spisu powszechnego). Równie problematyczny okazał się irlandzki spis powszechny, w którym miałam okazję brać udział razem z mężem.

Anegdotka o irlandzkim spisie powszechnym

Bo wiecie, mój mąż, Wojciech, to chłopiec miły, grzeczny, ale jednak z politechniki. A w życiu najbardziej ceni sobie precyzję. Pochylny wspólnie głowy, by uczcić tych bohaterów, którzy nigdy nie usłyszeli zakończenia historii opowiadanych przez mojego męża, bo po drodze umarli ze starości. Wojtuś wychodzi bowiem z założenia, że historie należy opowiadać dokładnie, precyzyjnie, a im więcej po drodze można wymienić szczegółów, tym lepiej. Wymienianie w jakiś dziwny sposób uspokaja umysł mojego męża. Na przykład ostatnio byliśmy w kinie i Wojtek opowiadał mi, w którym rzędzie nie mógłby siedzieć. Opowiadał tak:

W rzędzie A mógłbym siedzieć, w rzędzie B mógłbym siedzieć, w rzędzie C mógłbym siedzieć, w rzędzie D mógłbym siedzieć... (tu pomnę część tej historii, bo liczę, że załapałicie algorytm) w rzędzie J mógłbym siedzieć – kontynuuje Wojtek – ...w rzędzie K mógłbym siedzieć, o! I w rzędzie L już nie mógłbym siedzieć.

Oczywiście niektórzy mogliby pomyśleć, że istniał jakiś prostszy sposób przekazania mi tej informacji, ale to jacyś hipisi, anarchiści logicznego porządku wszechrzeczy. No i właśnie tacy sami anarchiści stworzyli irlandzki spis powszechny. Bo irlandzki cenzus ma tak jasno i logicznie skonstruowane pytania, że już w połowie człowiek ma ochotę zapisać się na jakieś zajęcia bardziej odpowiadające jego możliwościom intelektualnym, na przykład naukę wrzucania jabłka do wiadra albo robienia razem z szympansem Tadeuszem z pobliskiego zoo gryzaków z kamienia i kawałka sznurka. Jak przystało na człowieka, który uczy metodologii badań ankietowych, pomyliłam się tylko pięć razy, choć raz to raczej z własnej winy, bo kiedy kazali nam wpisać, ile osób mieszka u nas w mieszkaniu, to przychojraczyłam i postanowiłam policzyć w pamięci. I w ten właśnie sposób, jak dodałam jednego Wojtka i jedną Janinę, to mi wyszło trzy.

Serio, dobrze, że nie zarabiam na życie liczeniem, bo inaczej mogłoby być słabo.

Więc mój logiczny i precyzyjny mąż wypełniał spis i cierpiał. Cierpiał, bo pytania kwestionariusza w ogóle nie spełniały standardów dokładności Wojtka, no bo jeśli go pytają, jak długo z reguły jedzie do pracy, to co znaczy „z reguły”? Czy to znaczy, że wtedy, jak są korki, czy jak nie ma korków, czy ma policzyć średnią z całego miesiąca, czy medianę, a co, jeśli w ostatnim miesiącu przez trzy dni nie był w pracy, a raz pojechał później, bo nabrał ochoty na pączka i zajechał po niego na stację benzynową, ale w końcu go nie kupił, bo akurat nie było pączków, były tylko te takie ciasteczka z dżemem, ale to Wojtek nie przepada, bo ten dżem jakiś taki dziwny, więc nie kupił, no więc czy to się liczy do „z reguły”, czy to też ma dodać, czy nie dodać, wliczyć dzień z pączkiem czy nie wliczyć? Nic tu się nie układa w logiczną całość!

I im bardziej on ten cenzus wypełniał, tym bardziej ja miałam ochotę zmienić własny formularz, dokładniej rzecz ujmując – to pytanie na temat stanu cywilnego. No ale dobrze, Wojtek zaczął zmierzać do końca, i to w porę, bo ja już powoli dostawałam niedotlenienia mózgu od ciągłego powstrzymywania się od komentarzy, po chwili jednak mój mąż trafił na kolejną chłostę od życia. Ale co oni mnie tu pytają, jakimi językami mówię na co dzień – denerwował się – przecież mi tu nie starczy kratek na odpowiedź!

Skoro mu nie starczy kratek na odpowiedź, to iloma językami on mówi na co dzień? Polskim, angielskim, ludzi i aniołów?

Noooo... – zaczął wymieniać – Python, Java, C++...

Sami widzicie, podczas przygotowywania takich spisów powszechnych trzeba pomyśleć o szeregu istotnych rzeczy. Na przykład o ołówkach. Albo o tym, żeby spis był dostosowany do potrzeb informatyków.



A co, jeśli nie mamy wystarczającej liczby ołówków?

Na szczęście nie wszyscy badacze muszą przeprowadzać spis powszechny i ostrzyć ołówki dla całej swojej populacji. Mogą przeprowadzić badanie na mniejszej grupie dzięki pobraniu **PRÓBY BADANEJ**. Tak się dzieje najczęściej głównie z przyczyn praktycznych (ograniczone zasoby czasowe, finansowe, ludzkie...). Próba badana ma dokładnie takie same cechy jak populacja, to znaczy, że jeśli w naszej populacji mamy wszystkie dzieci chodzące do przedszkola w Warszawie, to w naszej próbie również będziemy mieć dzieci chodzące do przedszkola w Warszawie, tylko trochę mniejszą ich liczbę. A jaką dokładnie?

Wróćmy na chwilę do kwestii naszych dwóch koni i szopa pracza i zastanówmy się, ile tych koni i szopów musiałoby przebadać małżeństwo Warrenów, żeby ich badanie miało trochę więcej sensu. No mogliby na przykład skorzystać z wzoru na minimalną liczebność próby. Dobre wiadomości są takie, że nie będziemy się takimi wzorami zajmować, albowiem bardzo zależy mi na tym, byście dalej tę książkę czytali, a nie zrobili z niej origami... To, co wy musicie zapamiętać, zarówno jeśli chodzi o próbę badaną, jak i życie, to:

Czy wielkość ma znaczenie?

Wyobraźcie sobie, że chcecie poznać zwyczaje sportowe studentów Uniwersytetu Warszawskiego i w tym celu stajecie sobie nonszalancko nieopodal jakiejś studenckiej siłowni. Wiecie, jak gdyby nigdy nic, nikt nie

musi wiedzieć, że ostatnio biegaliście w 1995 roku, gdy rzucili drożdżówki ze śliwką do szkolnego sklepiku. No i stoicie sobie obok tej siłowni, a następnie pytacie każdą osobę, która z niej wychodzi, czy uprawia sport i jak często. Łatwo się domyślić, że na podstawie waszej ankiety okazałoby się, że nagle wszyscy są szalenie usportowieni, a liczba studentów kalających się wysiłkiem fizycznym sięga niemal stu procent (nie licząc tych, którzy poszli na siłownię tylko dlatego, że jest tam maszyna z batonikami¹⁹).

A co musielibyśmy zrobić, żeby uzyskać wynik bardziej odpowiadający rzeczywistości? Postawić jednego człowieka przed siłownią, a drugiego przed cukiernią? To wciąż nie rozwiązuje problemu. Z metodologicznego punktu widzenia najbliższe naszemu sercu jest rozwiązanie, w którym każda jednostka (w tym wypadku – student UW) ma takie samo prawdopodobieństwo wzięcia udziału w badaniu. Taką próbę nazywamy PRÓBĄ REPREZENTATYWNĄ, bo dobrze odzwierciedla całość populacji (wszystkich studentów UW). Taka próba jest pobierana w sposób losowy prosty lub z wykorzystaniem złożonych schematów losowego doboru próby. W przypadku naszego badania na studentach UW moglibyśmy na przykład wziąć listę wszystkich studentów rzeczoności uniwersytetu, wrzucić pozyskane nazwiska do pustego akwarium i wylosować 1000 osób, które wezmą udział w badaniu. Na uczelniach przeprowadza się takie losowanie raczej za pomocą komputerowego generatora liczb losowych, a nie losowania z akwarium, zwłaszcza że w naszym przypadku akwarium jest już zajęte przez karasie z poprzedniego badania.

Badania przeprowadzane na ludziach, którzy akurat wychodzą z siłowni albo mieszkają na tym samym piętrze lub też są naszymi znajomymi na Fejsie, więc akurat im wyświetli się nasza ankieta, to badania przeprowadzane na PRÓBIE NIEREPREZENTATYWNEJ (zwanej DOBOREM CELOWYM). Czyli takiej, w której pewne jednostki znajdują się w badaniu z większym prawdopodobieństwem niż inne. Wracając do przykładu ankiety spod siłowni – dla studentów, którzy akurat tego dnia udali się na trening, prawdopodobieństwo wzięcia udziału w badaniu wynosi jeden, a dla tych, którzy akurat plotą sobie naszyjnik z pączków w pobliskiej cukierni – zero. Łatwo się domyślić, dlaczego tak dobrana próba jest niereprezentatywna, a pozyskane wyniki mogą wprowadzać nas w błąd – ci, którzy znajdują się w naszym badaniu, mogą się systematycznie

różnić od reszty populacji. Taki błąd nazywamy BŁĘDEM DOBORU PRÓBY (ang. *sampling bias*).

Założmy, że chciałabym się dowiedzieć, jak moi studenci radzą sobie z materiałem, i w tym celu podczas piątkowego wykładu o 17.00 rozdaję im krótki test. Jaki będzie efekt? No efekt będzie taki, że w przyszłym tygodniu już nikt się na moim wykładzie nie pojawi. Oprócz mnie oczywiście. Znaczący, ja też kiedyś próbowałam nie przyjść, ale szef wyjaśnił mi, że nie można nie przychodzić na zajęcia, które się prowadzi. No i jeszcze może być tak, że wyniki będą znacznie lepsze, niż byłyby dla całej grupy, bo nie możemy wykluczyć, że studenci, którzy zjawiają się na wykładzie w piątek o 17.00, są bardziej zmotywowani i pilniejsi od pozostałych²⁰. Podobny problem występuje w wypadku umieszczania ankiet (na przykład do prac dyplomowych) na Facebooku – nasi znajomi nie są w żaden sposób reprezentatywni. Wynik może być szczególnie niemiarodajny, jeśli w owej ankiecie pytamy o kluczowe przekonania i opinie. A to ze względu na coś, co nazywamy HOMOFILIĄ W SIECIACH SPOŁECZNYCH. Co po ludzku oznacza tyle, że wykazujemy tendencję do przyjaźnienia się z ludźmi, którzy mają podobne upodobania, poglądy i opinie do naszych. Tak, to te słynne bańki, w których żyjemy.

Spektakularnym przykładem błędu doboru próby był sondaż dotyczący wyborów prezydenckich w USA w 1936 roku. Przed tymi wyborami „Literary Digest”, jedna z czołowych amerykańskich gazet, spytała ponad 2 miliony(!) Amerykanów o preferencje wyborcze i uzyskała wynik wskazujący na wygraną A. Landona (57% głosów). George Gallup, czyli absolutny psychofan sondaży, przepytiał 50 000 osób i na podstawie uzyskanych wyników uznał, że „Literary Digest” się myli, a wybory wygra F.D. Roosevelt. Mimo CZTERDZIESTOKROTNIĘ mniejszej próby Gallup miał rację – Roosevelt wygrał przewagą 63% głosów. To z kolei koronny dowód na to, że ja w sumie też miałam rację, mówiąc, że wielkość niekoniecznie ma aż tak duże znaczenie. Co innego jakość.

No, ale skąd aż taka pomyłka? Otóż gazeta przeprowadziła sondaż głównie wśród swoich czytelników. W dodatku był to sondaż telefoniczny. Pamiętajmy, że działo się to w latach 40., a więc telefon był dobrem luksusowym, a gazety czytali ludzie o konkretnym statusie społecznym. Krótko mówiąc: sondaż „Literary Digest” dotarł tylko do najbogatszych. I oni rzeczywiście częściej głosowali na Landona, również podczas prawdziwych wyborów. Niemniej ta próba totalnie wykluczyła resztę

populacji. Gallup lepiej przewidział wyniki wyborów, bo jego próba była losowa, a więc reprezentatywna dla całego społeczeństwa, nie zaś konkretnej grupy społecznej.

Zatem zawsze warto pamiętać o doborze próby – nieważne, czy o badaniu tylko czytamy i wysnuwamy na jego podstawie wnioski, czy też sami je przeprowadzamy. Bo wnioskowanie o populacji na podstawie źle dobranej próby to jak wysnuwanie wniosków na temat pogody w Polsce tylko na podstawie pomiaru temperatury w lipcu.

Co ma wspólnego foka z General Motors?

Istnieją takie badania, w których reprezentatywność próby nie ma większego znaczenia – są to badania jakościowe, czyli na przykład wywiady, badania terenowe czy obserwacja uczestnicząca. Niemniej zawsze musimy liczyć się z tym, że takie badanie jest niereprezentatywne, to znaczy, że nie możemy uogólniać jego wyników na całą populację. A więc z naszych ankiet spod siłowni nie możemy wysnuć wniosku, że większość studentów UW uprawia sport. Jedyne, co możemy powiedzieć na ich podstawie, to że wśród osób badanych zdecydowana większość podejmuje wysiłek fizyczny. Chociaż nie wiadomo po co i dlaczego.

Ale mimo wszystko obowiązkiem każdego badacza jest poinformowanie czytelników, w jaki sposób wybrał osoby do swojego badania. A czytelnicy – jeśli chcą być prymusami w metodologicznym przedszkolu – przy każdym badaniu opisywanym w gazecie, internecie czy czasopiśmie naukowym powinni zwracać uwagę na to, KOGO zbadano i ILE tego „kogoś” było.

Z tym że uwaga! Istnieje taki rodzaj badania, w którym liczba badanych osób, instytucji czy grup jest zawsze niewielka. To STUDIUM PRZYPADKU. Ta metoda badawcza polega na wykorzystywaniu wielu różnorodnych metod (na przykład wywiadu, testu psychologicznego, wywiadu z rodziną, obserwacji) po to, by jak najwięcej się dowiedzieć o konkretnym przypadku. Bohaterem studium przypadku niekoniecznie musi być człowiek. Tę metodę można wykorzystać również do opisu kraju, zjawiska społecznego czy sprawy sądowej. Na przykład Abadie i Gardeazabal (2003) chcieli zanalizować ekonomiczne koszty konfliktu wewnętrznego w danym kraju. W tym celu na swoje studium przypadku wybrali Kraj Basków, które od roku 1960 (z krótkim epizodem pokoju w latach 1998–1999) zмага się z aktami terroryzmu. Studia przypadków

na podobne tematy były przeprowadzane również w Irlandii Północnej (na przykład Cairns i Darby 1998, Fielding 2003) czy na Cyprze (Strong 1999). Znane są również studia przypadku jednej konkretnej firmy, na przykład Forda czy General Motors (Bayou i de Korvin 2008), albo jednego konkretnego gatunku zwierzęcia, na przykład wybranego gatunku foki (Hoelzel 1999).

Zapytacie: no ale jak to tak? Foka może być bohaterem studium przypadku, a szop i dwa konie nie? No nie. Bo jeśli zastanowimy się nad wspomnianym badaniem, to ten zwierzęcy gang w postaci dwóch koni i szopa pracza... nie miał żadnego metodologicznego uzasadnienia. Badacze chcieli przeprowadzić eksperyment, w którym uczyli zwierzęta wykonywania określonych czynności. Wybrali takie, a nie inne stworzenia, bo akurat takie żyły w okolicznej stajni. Dodatkowo ich liczba w żaden sposób nie pozwalała na wysnucie jakichkolwiek sensownych wniosków. Choćby nasze dwa konie w asyście jednego szopa poradziły sobie w tym eksperymencie najlepiej na świecie, to i tak nie świadczyłyby to w żaden sposób o możliwościach poznawczych tych gatunków. Wszak kto wie, może akurat mieliśmy do czynienia z przedstawicielem szopiej Mensy i dwójką najinteligentniejszych koni na świecie.

Obawiam się, że dla autorów tego badania nie ma żadnych okoliczności łagodzących. Będziemy ich odwiedzać w metodologicznym piekle, zupełnie tak samo jak ludzi, którzy opowiadają tę śmieszną anegdotkę, mówiącą, że hi, hi, jak wyjdiesz z psem na spacer, to hue, hue, hue, statystycznie macie po trzy nogi.

Ale to ogromna szkoda, bo osobiście uważam, że na uczelni mamy zdecydowanie za mało szopów praczy, a wszyscy moglibyśmy skorzystać z ich obecności. Na przykład w wypadku zepsucia się pralki wystarczyłoby usiąść na wykładzie koło kolegi szopa i przy okazji podsunąć mu miskę i kilka brudnych skarpet.

To, czy nasza próba musi być reprezentatywna, czy nie, jest ściśle związane z metodą badawczą, którą wybierzemy. O czym dowiecie się przy okazji odpowiedzi na pytanie, które zadawalibyście w licznych listach do mnie, gdybyście tylko je pisali. A mianowicie: skąd wiemy, że owca na ulicy rozpoznałaby prezydenta?

Skąd wiemy, że owca rozpoznałaby na ulicy prezydenta?

Mam pewnego przyjaciela, którego zresztą będziecie mieli okazję poznać w rozdziale *Dlaczego ufamy szarlatanom?*. Nazywa się Piotr Bucki i jest chodzącą dyskietką – wiem, że niektórzy z was zastanawiają się teraz, co to za śmieszne zwierzę „dyskietka” i gdzie je można spotkać²¹, więc już śpieszę z wyjaśnieniem. Młodszym czytelnikom tłumaczę, że dyskietka to taka bardzo stara wersja Dropboxa, z tym że lepsza, bo oprócz przechowywania na niej plików można też było przez nią oglądać zaćmienie słońca (a przez Dropboxa nie). Czyli mamy to – mój kolega jest takim chodzącym miejscem na dysku, który gromadzi tylko wartościowe informacje. Jeśli gdzieś przeprowadzono jakieś badanie naukowe, to on najpewniej o nim słyszał, a na dowód tego to ja mam taką anegdotkę, że kiedyś trafiłam na badanie, które wykazało, że owce rozpoznają twarze innych owiec (Peirce et al. 2001, Peirce et al. 2000, Kendrick et al 1995, Kendrick et al. 1996).

W sensie, że jeśli taki owiec Henryk zapozna się na pastwisku z owcą Barbarą, a trzy lata później znów się spotkają przy jakiejś koniczynie, to totalnie będą pamiętać, że już kiedyś sobie ściskali raciczki na przywitanie²². Strasznie mnie to podjarąło. Byłam w najwyższej fazie ekscytacji, kiedy zadzwonił do mnie Piotr, i pomyślałam, że to taki timing w sam raz, więc mu mówię, że on nie uwierzy, czego ja się dzisiaj dowiedziałam, takie badanie właśnie przeczytałam, które mnie szalenie podjarąło, albowiem okazało się, że owce totalnie rozpoznają twarze innych owiec. No, tak mu powiedziałam. A Piotr skomentował tę oszałamiającą anegdotkę w jedyny sensowny w takiej sytuacji sposób. Powiedział: dziki też.

Ale to nie wszystko! Chwilę później trafiłam na badanie, które wykazywało, że owce potrafią rozpoznawać również ludzkie twarze. Tak jest, jeśli taki Barack Obama zaprosi jakiegoś wełniastego przyjaciela na przyjęcie sylwestrowe, to trzy miesiące później, gdy spotkają się przypadkiem w supermarketowej alejce (miejmy nadzieję, że przy warzywach, a nie baraninie, bo byłaby niezręczność), owca go rozpozna. Mam nadzieję, że gdy to przeczytaliście, zareagowaliście dokładnie tak, jak powinniście reagować na wszelkie inne doniesienia naukowe i nagłówki

gazet informujące o oszałamiających wnioskach naukowych i statystykach, to jest: pomyśleliście sobie, że wszystko spoko, ale jak oni (naukowcy) to właściwie sprawdzili?

No serio. Logiczne przecież, że nikt nie przeprowadził ankiety wśród owiec, bo one mają takie nieporadne raciczki, że wieki całe by im zajęło zaznaczenie ołówkiem choćby jednej odpowiedzi. Pytanie o to, w jaki sposób badacz doszedł do określonych wniosków, jest szalenie istotne, a w metodologii nauk nazywa się DOBOREM METODY I TECHNIKI BADAWCZEJ. Metody badawcze w naukach społecznych to bardzo ciekawe zagadnienie. To jest tak, że kiedy ktoś na uczelnianej kawie rzuci luźne pytanie w stylu „Hej, co myślicie o badaniach etnograficznych?”, to chwilę później trzeba na tymże wydziale liczyć rannych i zabitych.

Kiedy literka poznaje cyferkę

ILOŚCIOWE METODY BADAWCZE to te umożliwiające nam znalezienie odpowiedzi na pytanie, które bardzo często słyszymy w cukierni. Nie, nie chodzi mi o: „Nie za dużo?!”, tylko o pytanie: „Ile?”. I nie, nie o liczbę kupionych pączków, ale o wszelkie interesujące nas kwestie. Jaki procent Polaków chodzi regularnie do kościoła? Ilu uczniów danej szkoły nie zdało matury? W jakich rejonach Puszczy Białowieskiej najczęściej występują gacki brunatne? Lub jeśli wciąż myślicie o pączkach z początku rozdziału: w które dni sprzedaje się najwięcej pączków? Jaki dochód przynoszą pączki?. Czyli rozumiecie, określamy tu parametry liczbowe badanego zjawiska lub obiektu (pączka), a następnie je analizujemy – najczęściej za pomocą analizy statystycznej lub modelowania ekonometrycznego.

Metody ilościowe to na przykład wszelkie badania ankietowe, kiedy to pytamy ludzi, czy zjedli kiedyś za dużo sernika, a potem liczymy, że 10% ankietowanych odpowiedziało TAK, 15% – NIE, a reszta nie zrozumiała, co to znaczy „za dużo sernika”. Wszelkie ankiety cieszą się dużą popularnością, zwłaszcza w naukach społecznych, o czym wie każdy, kto w najgorętszym okresie pisania prac dyplomowych (czyli jakieś dwa tygodnie po tym, kiedy powinny zostać złożone) nieroztropnie wszedł na Facebooka i został zaatakowany tysiącem linków do mniej lub bardziej przemyślanych kwestionariuszy²³.

Tymczasem przeprowadzenie dobrego badania ankietowego jest znacznie bardziej skomplikowane, niż mogłoby się wydawać.

Zaryzykowałabym nawet stwierdzenie, że znacznie bardziej skomplikowane niż zrobienie fikołka w przód. W tym ostatnim to ja jestem totalnym mistrzem – jak kiedyś potknęłam się na górze, to turlałam się w dół niczym najdoskonalsza na świecie rura PCV. Do tematu robienia ankiet tak pięknych, by wszystkie foki klaskały, a ludzie płakali ze wzruszenia, wrócimy jeszcze w rozdziale *Na imieniny lepiej kupić czekoladki czy kokainę?*.



Klasyczne badania eksperymentalne to również badania ilościowe. Na przykład gdybyśmy chcieli wiedzieć, czy jedzenie kalafiora jest dobre na mózg (no bo kalafior wygląda jak mózg), to moglibyśmy zaprojektować następujący eksperyment: stworzyć dwie grupy – eksperymentalną i kontrolną, i losowo przydzielić każdego z naszych uczestników do jednej z nich. Następnie przeprowadzić test na inteligencję, by sprawdzić, jak tam się ma IQ w obu grupach, a dalej to już z górki – zamykamy obie grupy w dwóch osobnych pokojach, jedną karmimy kalafiorem, drugą – wszystkim innym, tylko nie kalafiorem, a po określonym czasie ponownie przeprowadzamy ten sam test IQ co na początku naszego eksperymentu. Oczywiście po to, by sprawdzić, czy coś się zmieniło. Czy grupa, która uporczywie jadła kalafiora, naprawdę od tego zmądrzała i jest już w połowie drogi na wyspę geniuszy. Ewentualnie, by przetłumaczyć to wszystko z ludzkiego na naukowy, mogę powiedzieć, że eksperyment polega na manipulacji zmienną niezależną, stanowiącą przedmiot badania, aby sprawdzić, czy wpłynie to na zmienną zależną. I ja wiem, że to brzmi super, ale mam złe wiadomości – w praktyce w naukach społecznych rzadko kiedy mamy możliwość przeprowadzenia klasycznego eksperymentu. To znaczy zakładam, że znaleźlibyśmy kilkoro rodziców,

którzy zgodziliby się zamknąć swoje dzieci na kilka dni gdzieś z dala od cywilizacji, a przede wszystkim ich samych²⁴, ale jak niby przekonalibyśmy dzieci do jedzenia kalafiora?!

Bardzo często nie możemy przeprowadzić eksperymentu ze względów praktycznych lub etycznych, a także w pełni kontrolować wszystkich czynników, które mogą wpłynąć na jego wynik. Takie badanie, które nie spełnia wszystkich założeń, nazywamy quasi-eksperymentem²⁵. Eksperymenty tego typu spotykamy bardzo często w badaniach psychologicznych lub tych z dziedziny ekonomii behawioralnej. Jeśli wciąż nie podjaraliście się tym tematem, to może zainteresuje was to, że są i tacy, którzy za takie eksperymenty dostali Nagrodę Nobla²⁶. A to fajna nagroda, bo mama jest bardzo dumna.

Warto pamiętać o tym, że w ilościowej strategii badawczej lubimy duże, reprezentatywne próby, które pozwalają nam uogólnić wnioski na całą populację. Jednak ponownie podkreślę: istnieją dyscypliny takie jak psychologia czy ekonomia behawioralna, które nie są aż tak bardzo przywiązane do idei reprezentatywności próby, a my, rycerze błędu doboru próby, musimy sobie z tym jakoś poradzić.

BADANIA JAKOŚCIOWE często są przeprowadzane na małych, niereprezentatywnych grupach i to jest nawet OK. Bo też cel takich badań jest zupełnie inny. Metody jakościowe to takie, w których próbujemy dokonać jak najdokładniejszej, pogłębionej analizy interesującego nas zjawiska. Czyli w sumie jest to ten moment, w którym nauki społeczne czerpią z klasycznej antropologii. Nie skupiamy się tutaj na liczbach, ale na doświadczeniach jednostek i grup, ich historiach, uczuciach i przemyśleniach. Dane jakościowe mogą być zbierane poprzez wywiady, analizę dokumentów (pamiętniki, e-maile, wideo, fotografie) lub obserwację uczestniczącą, kiedy to badacz staje się członkiem badanej grupy i w ten sposób poznaje jej funkcjonowanie. W gruncie rzeczy z badaniami jakościowymi jest podobnie jak z badaniami ankietowymi – wydaje się, że to proste, bo przecież wszyscy potrafią zadać parę pytań i wyciągnąć wnioski. Ale słuchajcie, gdyby tak było, to wszystkie babcie uzyskiwałyby na wigilii tytuł doktora honoris causa wydziałów antropologii z całego świata.

Przeprowadzenie dobrych i rzetelnych badań jakościowych to wyzwanie. Czego nauczył nas swego czasu Bronisław Malinowski – który może i był niesłuchanie mądry, ale dziecka do chrztu bym mu nie dała potrzymać.

Czemu Bronisław Malinowski nie powinien trzymać dziecka do chrztu?

Musicie wiedzieć, że Bronisław Malinowski jest swego rodzaju legendą antropologii, bo jako pierwszy na świecie wyprowadził tę naukę w teren. To on przekonał wszystkich, że jeśli chcemy poznać jakąś kulturę, to musimy jak najwięcej zobaczyć, porozmawiać z członkami społeczności (w ich własnym języku!), uczestniczyć w ich życiu codziennym, a także wszelkich obrzędach i uroczystościach. W 1914 roku udało mu się zdobyć pieniądze na przeprowadzenie takich badań na Wyspach Trobrianda²⁷, co jest godne podziwu, biorąc pod uwagę, że większość naukowców to od 1914 roku próbuje zdobyć pieniądze na jakiegokolwiek badania, z czego niektórzy są już na 130 stronie wniosków grantowych (co nie jest znów aż takim złym wynikiem, dajcie spokój, to już prawie połowa!!!). To w czasie tej i kolejnych podróży napisał swoje najśłynniejsze dzieła *Argonauci zachodniego Pacyfiku* oraz *Życie seksualne dzikich*. Do obranego stylu badań terenowych podchodził na tyle poważnie, że po pewnym czasie społeczność, w której życiu uczestniczył, uczyniła go wodzem wioski. Czyli do tego momentu wszystko szło dobrze i Malinowski mógł trzymać do chrztu nie wiadomo jak dużą liczbę dzieci. Gdyby oczywiście takowy występował w badanych kulturach. To wszystko trwało jednak tylko do 1967 roku.

Wtedy to opublikowano osobiste dzienniki Malinowskiego z jego wypraw. Treść tych zapisków nie pozostawiała żadnych złudzeń. No raczej nie zaprosilibyśmy go na herbatkę do własnej cioci (cudzej w sumie też nie). Ten (jak dotychczas się wydawało) otwarty, tolerancyjny badacz, zakochany w idei zrozumienia drugiego człowieka i odmiennej kultury, okazał się megalomanem o skrajnie rasistowskich poglądach. Co nie przekreśla jego dorobku naukowego i wkładu w rozwój antropologii jako nauki, właściwie myślę, że jest wręcz odwrotnie – Malinowski nauczył nas wiele zarówno o tym, co powinniśmy, jak i o tym, czego nie powinniśmy robić w badaniach społecznych. W zasadzie bez znaczenia, czy jakościowych, czy ilościowych.

Dobra, czyli mamy to – wiemy już, że strategie badawcze dzielą się na dwie grupy, a te dwie grupy dzielą badaczy z taką precyzją, z jaką ja rozkrajam ostatni kawałek pizzy pomiędzy wszystkich zgromadzonych. 80% dla mnie i niech się dzieje wola nieba.

No myślę sobie, że gdyby kiedyś zorganizować bitwę na balony z wodą, to te dwie frakcje – zwolennicy metod jakościowych i metod ilościowych – stanęłyby w szranki. Oczywiście gdyby nie to, że taka bitwa to jednak trochę jakby sport, a jeśli chodzi o aktywność fizyczną, to typowy człowiek-akademik najczęściej myśli o kopaniu, a i to tylko w kalendarz i tylko w momencie otrzymania uwag od recenzenta.

Co metodologia ma wspólnego z majonezem?

Wigilia z miłośnikiem badań ilościowych i psychofanem jakościowych po przeciwnych stronach stołu byłaby koszmarem – ten pierwszy krzycałby, że co to za sztuka zadać trzy pytania na krzyż sąsiadce i jej dwóm koleżankom, to przecież żadna nauka. Każdy psychofan badań jakościowych odwdzięczyłby się zaś długim wywodem na temat tego, że człowiek to nie ciasteczko wycinane od szklanki, wszyscy jesteśmy różni, więc to świetnie, że chcecie nas zdegradować do rzędu uśrednionych cyferek, ale to ma tyle sensu co prasowanie majtek – można, ale po co? No i dajcie spokój, w takiej atmosferze to nie da się w spokoju podyskutować o paradygmatach badawczych, nie mówiąc o jedzeniu pierogów. Do bani taka wigilia.

Odwieczna wojna metod ilościowych i jakościowych o supremację przypomina niekończący się spór nauk ścisłych i humanistycznych, ewentualnie wojny miłośników majonezu o to, który jest smaczniejszy i dlaczego Kielecki. Podobnie jak w przypadku majonezu przychodzi z jedyną słuszną odpowiedzią: zarówno z dyscyplinami naukowymi, jak i metodami badawczymi jest jak z dziećmi – wszystkie trzeba kochać jednakowo²⁸. Choć oczywiście trudno w to uwierzyć, skoro przekonanie o antagonizmie nauk i wyższości jednej nad drugą panuje już od tylu lat. Żeby być dokładnym – od kilkuset. Przekonał się o tym bardzo boleśnie pewien brytyjski dziewiętnastowieczny poeta, Alfred Tennyson. Pomiziany przez natchnienie, napisał w 1842 wiersz *Wizja grzechu*, którego fragment brzmiał tak:

*Fill the cup, and fill the can:
Have a rouse before the morn:
Every moment dies a man,
Every moment one is born (...)*

Ten element poezji trafił w ręce matematyka Charlesa Babbage'a, który co prawda był szalenie mądrym człowiekiem (w końcu stworzył tablice logarytmiczne), ale na lekcjach literatury to najpewniej stał w kolejce do szkolnego sklepiku, albowiem po przeczytaniu Tennysonowskiego wiersza wystosował do autora list, w którym tłumaczył mu, że wszystko spoko, on wszystko rozumie, ale mieliśmy ostatnio wzrost populacji, więc ten wiersz to powinien raczej lecieć tak:

*Fill the cup, and fill the can:
Have a rouse before the morn:
Every moment dies a man,
Every moment ONE AND 1.167TH is born*

Żeby jednak poeta nie odniósł wrażenia, że jest jakimś tłumakiem, Babbage zapewnił go, że jest w pełni świadomy praw, którymi rządzi się poezja, on oczywiście totalnie rozumie, że należy uszanować prawa metrum, i dlatego jest w stanie również zaakceptować zaokrąglenie w postaci:

Co chwila umiera człowiek
I jeden koma dwa się rodzi.

I wiecie co? W drugim wydaniu poematu Tennyson zupełnie zignorował tę uwagę! Łączę się z nim w bólu, bo mnie, gdy opowiedziałam na imprezie żart o tym, „Wiecie, którego fizyka należy zawsze zapraszać na wesela? Heisenberga, bo zawsze będzie się bawić bombowo”, zignorowano dokładnie tak samo.

Osobiście jestem jednak daleka od myślenia, że matematycy całymi dniami rujnują życie jakimś poetom. Podobnie jak przedstawiciele nauk humanistycznych zajmują się w życiu czymś więcej niż tylko bieganiem po łące w lnianym worku na ziemniaki przepasanym witkami trzciny i rozmawianiem z motylami. Lubię przypominać, że wszystkie dyscypliny naukowe są tak samo ważne, co kiedyś wspaniale podsumowała reklama uniwersytetu w Utah, która głosiła: „Nauki przyrodnicze pozwolą Ci sklonować dinozaura. Nauki humanistyczne powiedzą Ci, dlaczego to nie najlepszy pomysł”.

Podobnie zresztą jest z technikami i metodami, którymi staramy się zbadać dane zjawisko – nie jest prawdą, że jedne są lepsze od drugich. W gruncie rzeczy wszyscy zapominają o tym, że to, jaką metodę wybierzemy, jest ściśle związane z naszym pytaniem badawczym. No bo dajmy na to te owce – jeśli chcemy wiedzieć, czy rozpoznają ludzkie twarze, to moglibyśmy zorganizować przyjęcie, na którym zapoznamy wszystkie owce z naszymi gośćmi, a po trzech miesiącach zorganizować kolejne przyjęcie, na które zaprosimy te same osoby i te same owce, i sprawdzić, czy owce wiedzą, z kim rozmawiają, czy przedstawiają się wszystkim od nowa. No dobra. Niby można by tak zrobić, ale umówmy się, że to nie byłoby najlepsze badanie na świecie, bo też dajcie spokój, komu by się chciało robić tyle krokietów z trawą i stokrotkami. Więc w przypadku owiec, ale też każdego innego badania, najpierw musimy się zastanowić, jak to zrobić najlepiej. Rzecz, o której swego czasu zapomniał pewien amerykański student.

Dzień dobry, czy chciałby pan porozmawiać o byciu gangsterem?

Sudhir Venkatesh był studentem socjologii na uniwersytecie w Chicago. W ramach badań do swojej pracy doktorskiej postanowił zapuścić się w miejsce, gdzie z reguły trudno spotkać studentów. Nie, nie do biblioteki. Chodzi mi o pustostany, które służyły za siedzibę członkom chicagowskich gangów. Sudhir był zdeterminowany, by poznać ich zwyczaje, rytuały i sposób działania. W tym celu udał się do miejsca, w którym urzędował gang The Black Kings. Niech nie zwiedzie was ta urocza nazwa, mówimy o jednym z największych gangów narkotykowych w Stanach Zjednoczonych. Kiedy Sudhir napotkał ludzi, którymi był zainteresowany, wręczył im pieczołowicie stworzony przez siebie... kwestionariusz ankiety. A wiecie, co zrobili ci mili członkowie gangu, gdy otrzymali te kilka pytań zapisanych na kartce?

Oczywiście z ochotą na nie odpowiedzieli.

Albo i nie. Przywiązali go do krzesła i trzymali uwięzionego przez 24 godziny. Następnie zaś niejaki J.T., lider gangu The Black Kings, udzielił Sudhirowi najważniejszej lekcji w życiu. Co ciekawe – dotyczyła ona metodologii. Powiedział mu mianowicie, że ej, człowieku, jeśli chcesz poznać i zrozumieć to, jak funkcjonujemy, to wyrzuć przez okno tę ankietę i po prostu poobserwuj, jak żyjemy, pogadaj i pobądź z nami, na własne

oczy zobacz, jak to funkcjonuje. Technika badawcza, którą zaproponowano Venkateshowi, nazywa się badaniem etnograficznym. Polega ona na jak najdłuższej obserwacji danej grupy w jej naturalnym środowisku i uczestniczeniu w jej życiu. Co Sudhir zresztą zrobił. Przez lata towarzyszył członkom gangu The Black Kings i zaprzyjaźnił się z jego liderem, który nawet w pewnym momencie pozwolił mu przejąć na jeden dzień swoje obowiązki w gangu. Z tych bardzo długich (siedmiu lat!) i bardzo dziwnych badań terenowych²⁹ powstał nie tylko doktorat, lecz także książka *Gang Leader for a day*, w której Venkatesh opisał swoje doświadczenia. Obecnie badacz jest profesorem socjologii na Uniwersytecie Columbia. I prawdopodobnie najbardziej klawym profesorem na świecie.

Oczywiście mamy w nauce również sytuacje odwrotne – kiedy to obserwacja uczestnicząca (czy jakakolwiek inna strategia jakościowa) nie jest najlepszym sposobem na rozwiązanie problemu. Na przykład w 1933 roku doktor Allan Walker Blair z uniwersytetu w Alabamie chciał zbadać, jaki wpływ na organizm człowieka ma ugryzienie pająka zwanego czarną wdową (dla biologicznych świrów podaję również łacińską nazwę – *Latrodectus mactans*). W tamtym czasie niewiele było jednak przypadków takiego pogryzienia, więc doktor Blair postanowił zrobić jedyną rzecz, która w tej sytuacji wydawała się mieć sens – dać się ugryźć jadowitemu pajakowi. Dajcie spokój, w gruncie rzeczy niewiele miał do stracenia, to znaczy owszem, życie, ale jeśli udałoby mu się przewyciężyć tę niewielką trudność, to dalej czekałby już na niego tytuł superbohatera nauk biologicznych, ewentualnie – jeśli możemy wyciągnąć jakiegokolwiek wnioski z filmów o Spider-Manie – po prostu superbohatera.

Trudno odmówić słuszności wynikom tych badań, choć konieczność tego poświęcenia wydaje się wątpliwa, zwłaszcza gdy przeczytamy oryginalny artykuł badacza, który tak podsumowuje wnioski z owego eksperymentu: „Jad wstrzyknięty przez ukąszenie dorosłego pająka *Latrodectus mactans* jest niebezpiecznie trujący dla człowieka”³⁰. Niemniej łatwo sobie wyobrazić, że owa jednoosobowa obserwacja uczestnicząca – nawet jeśli szalenie zaangażowana – nie mogłaby jeszcze stanowić podstawy do tego, by wyciągnąć jakiegokolwiek wnioski na temat wpływu trucizny na człowieka i objawów, które powoduje. Zresztą drugie ugryzienie nie mogło dać odpowiedzi na pytanie, czy wcześniejsze zetknięcie z trucizną wpływa na późniejszą odporność na ugryzienia. Przede wszystkim dlatego, że nigdy do niego nie doszło, bo Allana Blaira za bardzo bolało to pierwsze i zmienił

zdanie. Poza tym aby dojść do takich wniosków (lub je sfalsyfikować), i tak potrzebowalibyśmy szeregu wielofazowych badań klinicznych na większej grupie pacjentów³¹. A okazało się, że ludzi, którzy chcieliby zostać wielokrotnie ugryzieni przez pająki w celach naukowych, jest – co dziwne – niewielu. Ale ciekawe, czy ich liczba nie wzrosłaby, gdyby kandydatów zamknąć na 15 minut w pokoju, gdzie nie może wydarzyć się nic ciekawego, poza ukąszeniem przez pająka oczywiście.

Widzicie, to właśnie pytanie badawcze powinno decydować o doborze techniki badawczej. Byłabym ostrożna w badaniu niektórych grup społecznych lub tematów metodami ilościowymi. Z kolei wywiady lub grupy fokusowe nie pozwalają na wyciąganie wniosków na temat związków przyczynowo-skutkowych czy na generalizację wyników badań. Niemniej strategię badawczą można ze sobą mieszać po to, by zwiększać rzetelność swoich wniosków; by dokładniej zrozumieć naturę oraz poznać prawidłowości rządzące pewnym zjawiskiem. Na przykład moglibyśmy przeprowadzić ustandaryzowane testy wśród licealistów, by sprawdzić, jak sobie radzą ze stresem szkolnym i jaki jest ich stan psychiczny, a potem z częścią z nich przeprowadzić dodatkowo wywiady pogłębione, by dowiedzieć się więcej na temat ich osobistych doświadczeń, czyli tego, jak radzą sobie w szkole i w domu, poznać lepiej ich przyjaciół, rodzinę i grupę rówieśniczą. No chyba że naszą grupą badaną wcale nie będą licealiści, tylko owce, wtedy taka strategia badawcza byłaby absurdalna, bo wszyscy wiedzą, że jak owce są zestresowane, to po prostu liczą swoich znajomych skaczących przez płot.

Kiedy Harry poznał Sally, a prezydent owcę

Czy zastanawialiście się kiedyś, jakie filmy owce wybierają na pierwsze randki? No wiadomo, że komedie romantyczne, przecież nie ten okropny film z Anthonym Hopkinsem. Niemniej, o ile nie mamy żadnych wątpliwości, gdzie pan owiec mógłby zapoznać panią owcę, o tyle wcale nie jest dla nas oczywiste, gdzie owce miałyby spotkać prezydenta Obamę lub aktorkę Emmę Watson. A na tym opierało się badanie, o którym wspominaliśmy na początku rozdziału, mówiące o tym, że owce rozpoznają twarze nie tylko innych owiec, lecz także innych ludzi (choć my niekoniecznie potrafimy im się odwdzięczyć tym samym, chyba że wy umielibyście rozróżnić twarz jednej owcy od drugiej).

Wiemy już, że aby móc ocenić, na ile możemy ufać danemu badaniu, najpierw musimy przyjrzeć się temu, jak je przeprowadzono. W przypadku owiec wydaje się to jeszcze bardziej uzasadnione niż w przypadku ludzi, bo przecież żadna owca nie opowie nam, że ostatnie święta spędzała z owcą Henrykiem. Znajomością z prezydentem też się zresztą nie pochwali.



Eksperyment zaprojektowany przez naukowców uniwersytetu w Cambridge wyglądał tak: owcom pokazano serię zdjęć sławnych osób w losowej kolejności. Zdjęcia jednej osoby pojawiały się w puli kilkakrotnie, pod różnymi kątami (czyli na przykład Barack Obama en face, lewy profil Baracka Obamy, Barack Obama pozujący do średniowiecznego portretu). Owce oglądały sobie z uwagą wszystkie zdjęcia, a po jakimś czasie brały udział w kolejnej części eksperymentu, kiedy to owca wchodziła do pomieszczenia, gdzie pokazywano jej dwa zdjęcia – człowieka, którego już wcześniej widziała, i obrazek przedstawiający zupełnie przypadkową twarz. Gdy owcy udało się zgadnąć, która twarz jest jej znajoma, i pomiziać chrapkami³², to w nagrodę dostawała przysmak. Każda z ośmiu owiec zobaczyła serię dziesięciu par zdjęć i wśród tych dziesięciu par bezbłędnie wskazała osiem zdjęć, które już wcześniej widziała. Owcom szło trochę gorzej, gdy pokazywano im twarze pod różnymi kątami, ale wciąż w ponad połowie przypadków były w stanie rozpoznać swoich ludzkich przyjaciół. I teraz ciekawostka: pamiętały te twarze przez kolejne dwa lata³³!

Gdy zaznajomimy się bliżej z tym eksperymentem, to po pierwsze, znacznie łatwiej nam będzie usadzić znajomą owcę przy wigilijnym stole, bo będziemy mieć większą pewność, kogo z naszych gości już zna. Po drugie, łatwiej nam będzie ocenić, czy dana metoda badawcza była

odpowiednio dopasowana do pytania badawczego, a także do obiektów badawczych (czyli owiec). Gdyby mi ktoś powiedział, że przeprowadził wśród owiec serię wywiadów pogłębionych, to byłabym sceptyczna (dajcie spokój, one ciągle żują trawę, przy odpowiadaniu na pytania musiałyby strasznie seplenić). Taki eksperyment natomiast przekonuje mnie trochę bardziej. Podobnie jak metoda etnograficzna wydaje się bardziej efektywna do zbadania miejskich gangów niż ankieta. I jak widzicie – i jedna, i druga metoda może być przydatna. Wszystko zależy od tego, co lub/i kogo potrzebujemy zbadać.

Aha, i jeszcze jedno. Nie jest tak, że naukowcy tarmosili zdjęcie prezydenta Baracka Obamy owczymi chrapkami ot tak, dla zabawy. Eksperyment ten został zaprojektowany przez grupę neurobiologów po to, by ocenić, czy owcze zdolności rozpoznawania twarzy są porównywalne z innymi naczelnymi, w tym z człowiekiem. To odkrycie miało ogromne znaczenie dla dalszych badań nad chorobami neurodegradacyjnymi u ludzi. W gruncie rzeczy ten pozornie śmieszny eksperyment był tylko wstępem do szerszych badań dotyczących Alzheimerera. Ale czego miało dowieść badanie, w którym odkryto, że kurczaki chętniej dziobią w zdjęcia ładnych ludzi, to już naprawdę nie wiem³⁴.

Niemniej i jedno, i drugie badanie nie miałyby sensu, gdybyśmy otrzymanych wyników nie zanalizowali. Metod analizy, zarówno badań jakościowych, jak i ilościowych, jest mnóstwo. Analiza danych jakościowych obejmuje na przykład analizę treści, analizę dyskursu czy analizę struktur narracyjnych. W przypadku danych ilościowych mamy do wyboru szereg testów i modeli statystycznych, choć z reguły zaczynamy od czegoś znacznie bardziej podstawowego – statystyk opisowych. Wśród których mamy absolutną gwiazdę socjometryczną wszelkich dyskusji o sensowności nauki, statystyki i w ogóle liczenia, takiego wujka Wiesia, bez którego nie odbędzie się żadne wesele – średnią arytmetyczną.

Czy kiedy wychodzimy z psem na spacer, to średnio mamy po trzy nogi?

Pierwsze zajęcia w roku akademickim zawsze były dla mnie trudne, a to głównie za sprawą miejsca, gdzie odbywała się owa pedagogiczna prywatka, rewia najdoskonalszych popisów w zakresie szkolnictwa wyższego i prywatny benefis mojego edukacyjnego geniuszu, czyli że za sprawą Irlandii. Takie zajęcia trzeba było bowiem zaczynać od sprawdzenia obecności, a to zawsze trauma. Z jakiegoś powodu Amnesty International nigdy nie chciało przyznać, że nazywanie dzieci Aoibheann, Caoilfhionn czy Feidhlimidh jest wyrazem prowadzonej przez Irlandczyków polityki nienawiści wobec innych narodowości. Choć tak po prawdzie to to nie jest jakiś ogromny kłopot, bo czytanie irlandzkich imion wymaga jedynie odrobiny intuicji. Na przykład takie przyjemne imię jak Méadhbhh czyta się dokładnie tak, jak się pisze, tj. Maeve. No bo w sumie dlaczego nie. O tym fakcie poinformował mnie zresztą kolega Toirdhealbhaich, czyli kolega o imieniu, które ma piękne, historyczne pochodzenie – to znaczy nie mam żadnych wątpliwości, że po celtycku Toirdhealbhaich znaczy „dręczący naród polski”.

No to zaczynam! Zaczynam od zapoznania się z moimi nowymi studentami, proszę bardzo, niech każdy się przedstawi, a ja potem powtórzę i zapamiętam wszystkie imiona, i potem będziemy już mogli wyruszyć we wspaniałą, intelektualną podróż po oceanie nauk społecznych. I proponuję zacząć od koleżanki z prawej, na co koleżanka z prawej uśmiecha się do mnie szeroko i mówi głośno i z radością, jakby wcale nie chciała zaraz wbić fonetycznego noża w moje polskie serduszko: *My name is Sadhoh*. Sadhoh, cóż za piękne irlandzkie imię, myślę sobie, takie imię w sam raz, jeśli chcesz, żeby pewien nauczyciel z Polski już nigdy nie zadał ci żadnego pytania. No trudno, może jej koleżankę uda mi się kiedyś o coś spytać, proszę bardzo, zapraszam do przedstawienia się koleżankę obok koleżanki Sadhoh, a koleżanka obok mówi: *My name is Ailbhe*, a ja myślę sobie, że *of course it is*. Koło koleżanki Ailbhe siedzi student płci męskiej, włosy blond, ułożone w loki, oczy niebieskie, policzki okrągłe i rumiane, oto jest wcielenie niewinności, to musi być człowiek o imieniu z doskonale wyważoną liczbą samo- i spółgłosek, nie mam

żadnych preferencji, ale najlepiej by było, gdyby nazywał się Krzysztof. Więc pytam blond chłopca, jak ma na imię, pytam pełna nadziei i dziecięcej ufności, ja jestem jak taki buldog Rafał, który liże wam palce, bo ma nadzieję, że za plecami trzymacie kawałek pasztetowej. I ja, niczym ten ufny buldog, pytam go, jak ma na imię, a on mi na to mówi: *My name is Kyoosti*. Czaicie, Kyoosti. Kyoosti! Tak powiedział i tego już było za wiele. Więc tłumaczę moim nowym studentom, że ha, ha, ha, wszyscy się pośmialiśmy, ho, ho, ho, wyborna to komedia, ale dość tego, niech będą poważnymi i dorosłymi ludźmi, no bo po pierwsze, ktoś w tej sali musi być, a po drugie, to nie jest miłe, bardzo to nieeleganckie, otóż nie wolno im wymyślać szalonych imion tylko po to, żeby zbić z tropu mnie, człowieka-Polaka. A jak już muszą wymyślać, to niech chociaż zachowają minimum wiarygodności, niech przynajmniej używają imion, które mogłyby uchodzić za prawdziwe, a nie że wymyślają słowa, które co najwyżej mogą symulować dźwięk wydawany przez strusie w czasie godów. No, tak im powiedziałam. Na co blond chłopiec podniósł niepewnie rękę i cicho, jakby ze wstydem, oświadczył: *But that's my real name, I am from Finland*.

Więc od tamtego czasu z niecierpliwością oczekuję od mojego szefa zaproszenia na *diversity training*, niemniej w tamtym momencie postanowiłam zatrzymać te płynące wewnątrz Janiny wodospady łez utraconej godności, odłożyć na bok listę obecności i przejść do głównego tematu zajęć, który tego dnia – jako i każdego dnia – miał być szalenie ekscytujący. Wyjaśniłam moim studentom, co będziemy robić, a podekscytowana przy tym byłam jak ten sam struś, co to mu się powiodły te gody. I tłumaczyłam im entuzjastycznie, że teraz obliczymy i zanalizujemy sobie jakąś korelację, narysujemy jakiś miły wykres rozrzutu i jeszcze, żeby było sympatyczniej, weselej, to każdy będzie mógł wybrać swoje własne zmienne, tak jest, WYBRAĆ SWOJE WŁASNE ZMIENNE!!! *Oh, that's so exciting!* – chwale sama siebie, swój własny geniusz klepię z uznaniem po plecach, studentów zaś pytam, czy mają jakieś pytania.

I jest, ktoś się zgłasza! Ktoś ma pytanie! Oto mój pierwszy prymus nowego semestru, patrzcie, jak się rwie do odpowiedzi, jaka jest w nim pasja i ochota, by jak najszybciej wdać się ze mną w intelektualną potyczkę, napoić swój chłonny jak gąbka umysł życiodajnym płynem statystycznej wiedzy, a ja przy tym wszystkim muszę udawać całkowity spokój i zachowywać się jak gdyby nigdy nic, jakby ta ręka w górze to była

dla mnie najbardziej naturalna sytuacja na świecie i wcale nie jarała mnie jak truskawka żółwia. Tak, przypomnijmy sobie jeszcze raz tę wzruszającą scenę: korelacje, wykresy rozrzutu i *so exciting!*, a wtedy zgłasza się miły młody człowiek z ostatniej ławki i mówi, że ma pytanie. *I have a question* – mówi – *do you know the meaning of the word exciting?*

Przyznam wam, że trochę się zdenerwowałam, zupełnie jak mój szef wtedy, gdy powiedziałam mu, że niestety nie mogę już dłużej uczyć studentów położnictwa, to znaczy – bardzo bym chciała oprócz tego, że nie chcę. Mam bowiem wrażenie, że ci studenci tak dziwnie na mnie patrzą, jak takie zdziwione lemury patrzą, i to tylko dlatego, że na ostatnich zajęciach przez przypadek trzasnęłam w lalkę noworodka drzwiami, a potem jeszcze na nią przypadkiem nadepnęłam, a potem, kiedy próbowałam ją podnieść, to znowu wypadła mi z rąk i wtedy ja się tak zaaferowałam, że ją kopnęłam. Zdenerwowałam się zaś z dwóch powodów: po pierwsze dlatego, że ja w ogóle nie lubię ludzi takich jak ten student, takich ludzi, którym miast ludzkich, ciepłych uczuć i zwykłej życzliwości przewodzą szyderstwo, cynizm i głupie teksty. Nie przepadam za takimi ludźmi, bo ja generalnie bardzo źle znoszę konkurencję. Po drugie, moje niezadowolenie wynikało z tego, że wtedy jeszcze nie wiedziałam, że to nie jest najgorsze pytanie, jakie przyjdzie mi usłyszeć na mojej zawodowej, statystycznej drodze. A przede wszystkim – nie to, które przyjdzie mi usłyszeć najczęściej. Codziennie, a czasem i kilka razy dziennie. Takie pytanie babcia, które wraca jak bumerang, by upewnić się, że wciąż nie chcecie kluseczek, mimo iż minęło już 30 sekund od chwili, gdy ostatni raz odmówiliście. A czym jest ten statystyczny odpowiednik pytania babci, wyjaśni nam koncepcja co prawda teoretyczna, lecz szalenie wymowna, czyli ptyś z kremem o zapachu nieprzystającym.

Metaforyczny ptyś zalatujący kiełbasą

Musicie wiedzieć, że w cukierni wszystkich dyscyplin naukowych statystyka bywa takim spektakularnym ptyśkiem z kremem, który jest obietnicą najśłodszych partytur wygrywanych na kubkach smakowych, a tak naprawdę został zrobiony ze śmietany w proszku i jeszcze trzymano go w lodówce obok kiełbasy, więc cały krem smakuje jak stęchła jałowcowa. I widzicie, w takim przypadku nie jest winą ptyśki, że został

zrobiony z fatalnych składników, i nikomu nie przysłoby do głowy, by owego ptysia zgłosić do prokuratury.

Idąc tym tropem, trudno oczekiwać, że jak sobie zwerbujemy trzy osoby spod sklepu monopolowego, najlepiej po godzinie osiemnastej i po półlitrowce (niejednej), i każemy im trzasnąć sudoku, to z takich danych uzyskamy rzetelne wnioski dotyczące zdolności poznawczych całej populacji. Kiedyś jeszcze tłumaczyłam moim studentom, że ta zasada nazywa się *garbage in – garbage out*. Że jeśli będą analizować liche dane, takie numeryczne śmieci, to wnioski wyciągnięte na tej podstawie będą podobnej jakości. Niemniej szybko zaprzestałam używania tej metafory, bo okazało się, że jak mówię o *garbage out*, to im się natychmiast włącza zespół stresu pourazowego, na pamiątkę tego, że w zeszłe święta matka kazała im wynieść śmieci.

Istnieje również inny sposób mówienia o tym problemie, szalenie rozpowszechniony wśród członków pewnej grupy, która do statystyki ma podobne podejście jak ja do ćwiczeń fizycznych, to znaczy: nie lubię, nie znam się, robię wszystko, by unikać, ale jeśli trzeba sprawę skrytykować, to pierwsza będę stała z transparentem pod siłownią, krzycząc: „Wolność dla bicepsów o konsystencji maślanych bułeczek!!!” oraz: „Niech żyją nam mięśnie brzucha kruche jak dobrze wypieczona beza!!!”. Podobnie zachowuje się pewne internetowe plemię, które możemy rozpoznać po bardzo specyficznym zawołaniu, używanym namiętnie we wszelkich dyskusjach dotyczących statystyk, badań lub wykresów. Owo zawołanie wymaga zaawansowanej wiedzy matematycznej (dodawanie, dzielenie), zoologicznej (znajomość anatomii psa), a także medycznej, ze szczególnym wskazaniem ortopedii (rozeznanie w ludzkich kończynach dolnych).

I to by się w sumie zgadzało, oprócz tego, że się nie zgadza.

Zanim jednak wyjaśnię, na czym polega błąd, pragnę zaznaczyć, że w historii nauki zdarzało się już wcześniej takie połączenie statystyki i kynologii. Bo nie wiem, czy wiecie, że Francis Galton, miły naukowiec z przełomu XIX i XX wieku, który lubował się w metodach statystycznych i był psychofanem badań kwestionariuszowych, to owszem, jako pierwszy stworzył teoretyczną koncepcję korelacji i regresji do średniej, ale oprócz tego zawdzięczamy mu wynalezienie gwizdka dla psów. Ja w ogóle lubię historie o ludziach, którzy wynaleźli jakiś pierwiastek, a jeszcze przy okazji operę i kilka sonetów trzynastozgłoskowcem trzasnęli. Ja zawsze sobie

wtedy myślę, że spoko, dobrze, że ja w tym czasie zdążyłam jedynie zrobić test, jakim rodzajem pizzy jestem.

I jeszcze mi wyszło, że hawajską.

Ale słuchajcie, gdybyśmy sobie teraz skorzystali z tego wynalazku Galtona i zawołali do siebie jakiegoś psa, i policzyli, ile on ma łap, a potem policzyli, ile my mamy nóg, a następnie obie wartości zsumowali, a potem podzielili przez liczbę osobników, to wynikiem naszych obliczeń byłyby: po pierwsze, liczba trzy, a po drugie, okropne upokorzenie przed naszym osobistym dalmatyńczykiem. Albowiem w tym krótkim rachunku matematycznym popełnilibyśmy cztery różne błędy: pierwszy dotyczący POPULACJI, drugi dotyczący SKALI POMIARU, trzeci dotyczący MIARY TENDENCJI CENTRALNEJ (odpowiadającej owej skali) i czwarty dotyczący WNIOSKOWANIA STATYSTYCZNEGO. Podziękujmy wszystkim psom tego świata. Patrzcie, ile się dzięki nim nauczymy!



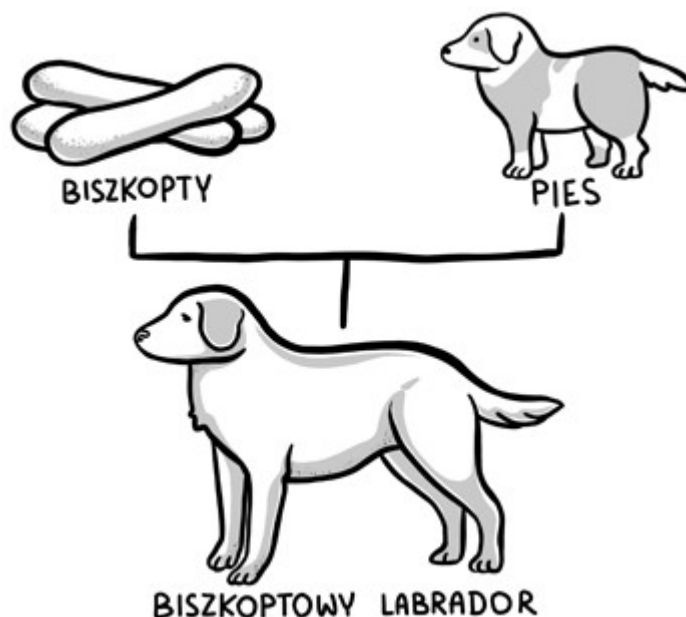
Czy jasne labradory pochodzą od biszkoptów?

Pamiętacie, w poprzednich rozdziałach nauczyliśmy się, że gdy przeprowadzamy badanie na jakikolwiek temat, to musimy wybrać populację badaną, czyli grupę (osób, zwierząt, krajów, instytucji...), która w tym naszym naukowym rendez-vous weźmie udział. Mogą do niej należeć określony gatunek nietoperza albo też wszystkie nietoperze zamieszkujące konkretny obszar (na przykład Puszcę Białowieską czy strych stryjenki). Wróćmy na chwilę do wspomnianego wcześniej badania o kurczakach, które lubią ładnych ludzi. Gdy mili badacze chcieli swego czasu sprawdzić, czy jest sens tworzyć Tindera dla kurczaków³⁵, potrzebowali w tym celu pobrać próbę z populacji ludzi i z populacji

kurczaków (Ghirlanda et al. 2002). Podobnie na próbie kurczaków pracowali naukowcy z Kanady (Newberry et al. 1985), którzy chcieli sprawdzić, jaki jest stosunek tych miłych pierzastych stworzeń do grzęd wazących³⁶. Waszą populację mogą stanowić nawet kraje biorące udział w konkursie Eurowizji, zupełnie tak samo, jak było w badaniu Ari Kokko i Patrika Tingvalla (2012). Jak widzicie, populacją, z której pobieramy próbę³⁷, mogą być i ludzie, i kurczaki, a jeśli jest potrzeba, to nawet rzeki (Jeffers 1998) czy czekolada (Padley i Timms 1980). A wiecie, czemu to ma znaczenie dla naszych kynologiczno-ortopedycznych rozważań?

No bo słuchajcie, ja mam dla was taką szokującą informację, że my to nie jesteśmy jakoś wybitnie spokrewnieni z psami, my jesteśmy dwoma odrębnymi gatunkami, i ja rozumiem, że to szalenie rozczarowujące, ale pomyślcie sobie, że jest jeszcze gorzej – słyszałam, że podobno jasne labradory wcale nie mają wspólnego drzewa genealogicznego z biszkoptowymi ciastkami (?!).

Czy możemy sobie zdefiniować populację jako „psy i ludzie” i na jej podstawie liczyć różne rzeczy? No możemy, tylko wtedy musimy się liczyć z tym, że nagle okaże się, że jeśli chodzi o umiejętności, to nasza populacja jest przeciętna zarówno w aportowaniu patyków, jak i w rozwiązywaniu sudoku. I że średnio mamy po ileś tam łąt, a jeśli chodzi o spędzanie czasu wolnego, to większość naszej populacji najbardziej lubi tarzać się w zwłokach zająca znalezionych w lesie.



Z liczeniem różnych rzeczy na podstawie tak zdefiniowanej populacji jest trochę jak z bieganiem dla przyjemności – można, ale po co?

Takie wnioski są po prostu wbrew zdrowemu rozsądkowi. Anglicy mówią na to „porównywanie jabłek i pomarańczy” (*comparing apples to oranges*), co oznacza absurdalne i nieprawidłowe porównywanie dwóch zupełnie odrębnych grup. W naszym wypadku – ludzi i ich włochatych przyjaciół o trójkątnych noskach.

A to jeszcze nie wszystko! Bo nawet jeśli ktoś by się uparł, że chce iść do piekła i tak właśnie sobie tę populację zdefiniuje, że będą w niej ludzie i labradory, to wciąż musi wziąć pod uwagę jeszcze jedną pralinę ze statystycznego pudełka rozkoszy...

Jak bardzo jesteś jamnikiem?

Wiecie, co jeszcze jest złego w stwierdzeniu o spacerze, psie i trzech nogach, oprócz tego, że wszystko? To, że w tym stwierdzeniu słowo „statystycznie” jest tożsame ze „średnio”, w domyśle zaś – ze średnią arytmetyczną (chyba że mowa o tym badaniu z grzędami i kurczakami, bo tam to najpewniej liczyli średnią ważoną, ha, ha, czaicie? WAŻONĄ³⁸). A statystyka i średnia mają się do siebie jak kwadrat do prostokąta – każde liczenie średniej jest działaniem statystycznym, ale nie każde działanie statystyczne musi być liczeniem średniej. Jest nawet gorzej – nie zawsze tę średnią w ogóle powinniśmy liczyć.

A to dlatego, że mamy w statystyce cztery różne skale pomiarowe, które opisują, w jaki sposób mierzymy różne rzeczy. A takich rzeczy do zmierzenia to mamy bez liku – spójrzmy na przykład na te warszawskie przedszkolaki. Możemy pomierzyć ich wiek, wzrost, masę ciała, płeć, ile mają pluszaków, ale również to, czy boją się pajaków i czy lubią budyń (choć to ostatnie byłoby bez sensu, no bo kto nie lubi?). Te wszystkie rzeczy będziemy mierzyć na jednej z czterech skal. Zanim przejdziemy do skal, odpowiedzmy sobie na podstawowe pytanie: czym jest pomiar? Pomiar to inaczej kwantyfikacja, czyli przypisanie wartości liczbowych do obserwacji. Czasem przypisywanie liczb jest arbitralne, definiowane przez badacza, dlatego nie zawsze sensowne będzie wykonywanie działań arytmetycznych na takich liczbach.

SKALA NOMINALNA to skala, której wartości są kategoriami bez żadnego oczywistego uporządkowania. O dwóch obserwacjach na skali

nominalnej możemy tylko powiedzieć, czy są takie same, czy różne. Przykładami zmiennej nominalnej są kolor oczu, zawód, płeć lub w – przypadku psów – rasa czy umaszczenie. Możemy policzyć, że w naszym psim przedszkolu mamy dwa jamniki, trzy labradory i pięć najpiękniejszych na świecie kundelków, niemniej nie możemy tych trzech kategorii w żaden sposób uszeregować na skali bardziej pies – mniej pies. Byłoby to matematycznie bezsensowne, a przy tym szalenie nieuprzejme względem kudłatych cudeniak. Nie miałyby też sensu mnożenie jamników przez labradory lub wyliczanie z nich średniej.

SKALA PORZĄDKOWA to wartości, które są kategoriami dającymi się uporządkować. Wiemy, czy dwie obserwacje są takie same, czy różne. A jeśli są różne, to która jest bardziej, ale nie o ile bardziej. Nie da się jednak zinterpretować odległości między poszczególnymi kategoriami. Przykładem zmiennej porządkowej jest wykształcenie lub w przypadku psów – wielkość mierzona na skali: mały pies – średni pies – duży pies. I takie kategorie można już sobie uporządkować, na przykład powiedzieć, że na edukacyjnej drabinie wykształcenie podstawowe znajduje się niżej niż wykształcenie wyższe lub że mały pies jest mniejszy od dużego psa, nie wspominając już o tym dogu niemieckim z za płotu, którego psia mama najpewniej miała romans z koniem.

Na zmiennych porządkowych, zupełnie jak w przypadku zmiennych nominalnych, wciąż nie możemy przeprowadzić żadnych poważnych operacji matematycznych. No chyba że ktoś naprawdę twierdzi, że ma jakikolwiek sens policzenie sobie średniej z wszystkich stanowisk i odkrycie, że w sumie to wszyscy są w 0,67 prezesem. No słuchajcie, jakbyście przedstawili te wyliczenia w fokarium, to obawiam się, że żadna foka by nie zaklaskała.

SKALA PRZEDZIAŁOWA (interwałowa) to skala, w której różnice między wartościami mają określoną interpretację. Innymi słowy, wiemy, czy dwie obserwacje są identyczne, czy różne, wiemy, która jest bardziej i o ile, ale nie znamy ich stosunku wielkości. Bo taka skala ma umownie przyjęty punkt zero (co oznacza, że może przyjmować wartości ujemne). Przykłady zmiennej przedziałowej to temperatura w stopniach Celsjusza lub daty urodzenia. Ja urodziłam się 11 kwietnia (ewentualnie – jak napisałam ostatnio w oficjalnym formularzu – 111 kwietnia), mam kolegę z podstawówki o imieniu Wojtek, który urodził się 11 lutego. Nasze daty urodzenia dzieli 60 dni w roku kalendarzowym. Wiem, kto pierwszy będzie

obchodził urodziny w danym roku oraz jeśli dziś są urodziny Wojtka, to za ile dni będą moje. Nie mogę jednak nic powiedzieć o stosunku wielkości między wiekiem moim i Wojtka, bo jest on zależny od tego, gdzie ustale umowne zero (rok urodzenia).

SKALA ILORAZOWA ma wszystkie właściwości trzech poprzednich skal, a dodatkowo ma jasno określony punkt zerowy, który świadczy o tym, że dana zmienna nie występuje. Przykładami zmiennej ilorazowej są wzrost, wiek (mierzony w latach) czy masa. Czyli wiemy, że dwie obserwacje są identyczne lub różne. A jeśli różne, to która jest bardziej, o ile i jaki jest stosunek ich wielkości. Na takich zmiennych możemy dokonywać wszelkich operacji matematycznych, możemy na przykład powiedzieć, że ktoś jest dwa razy cięższy od kogoś innego. Chociaż nie powinniśmy, bo to jednak nigdy nie jest miłe.

W uproszczeniu: lubimy mówić, że skale nominalna i porządkowa to skale jakościowe, a pozostałe dwie – ilościowe. W ten sposób łatwo zapamiętać, że o ile w przypadku skal jakościowych możemy sobie policzyć częstotliwość występowania (trzy jamniki, siedem kundelków) czy procenty (30% jamników, 70% kundelków), o tyle wszelkie poważniejsze operacje matematyczne (wliczając w to liczenie średniej) to rozkosz zarezerwowana tylko dla zmiennych ilościowych. Możemy policzyć średnią masę ciała, średni wzrost, średnią temperaturę.

W naszej grupie jamników i kundelków możemy policzyć również średnią długość psa, niemniej średniej bycia jamnikiem – już nie.

I teraz uwaga, proszę o przygaszenie świateł i cichy dźwięk werbli. Słuchajcie, liczba nóg to skala porządkowa, choć może wam się to wydawać sprzeczne z intuicją. Długość nóg, na przykład mierzona w centymetrach, to oczywiście skala ilościowa, która pozwala nam na wiele – możemy sobie wszystkich zmierzyć i uznać, że przeciętna długość nóg w naszej grupie to 56 centymetrów, możemy też policzyć, że ktoś ma dwa razy dłuższe nogi od kogoś innego.

Niemniej liczba nóg to skala porządkowa³⁹, nie ilościowa, nawet jeśli jest wyrażona liczbami. To się niestety zdarza i może wprowadzać w błąd, bo jest wbrew intuicji. Nawet jeśli wartości zmiennej nominalnej lub porządkowej są wyrażane liczbowo, to liczby te są tylko umownymi identyfikatorami, nie powinno się więc na nich wykonywać żadnych działań arytmetycznych.

W naturze występują jedna noga, dwie nogi, trzy nogi, cztery nogi, ale raczej nie powiemy do kogoś: „Hej, pamiętasz Mariana? To ten, co ma 1,76 nogi!” (niby możemy, ale wtedy trzeba się liczyć z tym, że już nigdy nie zaproszą nas na żadną imprezę). Możemy więc stwierdzić, że cztery nogi to więcej niż dwie nogi, ale liczenie średniej jest tutaj zwyczajnie nieprawidłowe. A żeby wyjaśnić dlaczego, musimy wprowadzić nowe pojęcie – MIARY TENDENCJI CENTRALNEJ. Zanim to zrobimy, odpowiedzmy sobie, na jakie pytanie odpowiada miara tendencji centralnej. Otóż miara tendencji centralnej odpowiada na pytanie, co jest najbardziej typowe. A skoro już wiemy, że mamy różne skale pomiarowe, które różnią się tym, jakie operacje matematyczne można na nich wykonywać, to mamy do nich pasujące bardziej lub mniej miary tendencji centralnej.

Lizak Chupa Chups świata statystyki

Tendencja centralna jest trochę takim lizakiem Chupa Chups świata statystyki. Słodka to rozkosz, ale zapakowana w taki sposób, że aby się do tego lizaka dostać, trzeba wezwać ślusarza, operatora młota pneumatycznego i dwa zastępy straży pożarnej.

Podobnie tendencja centralna – to bardzo proste pojęcie, opakowane jednak w podstępnie trudną nazwę. Sami zobaczcie: odnalezienie miar tendencji centralnej polega na określeniu wartości przeciętnej dla interesującej nas zmiennej lub grupy wyników. Tego interesującego nas środka. Możemy to zrobić na trzy sposoby, bo wyróżniamy trzy podstawowe miary tendencji centralnej:

Dominanta (wartość modalna, moda)

To wartość najczęściej występująca w danym zbiorze. W naszym zbiorze trzech jamników i siedem kundelków dominantą jest kategoria „kundelek”, bo jest to wartość najczęściej występująca.

Mediana

To wartość środkowa, powyżej i poniżej której znajduje się jednakowa liczba obserwacji. By wyłonić medianę, najpierw musimy uporządkować

nasz szereg. Czyli jeśli mamy w domu pięć małych psów i dwa duże, to nasz uporządkowany szereg psów wygląda tak:

mały pies – mały pies – mały pies – mały pies – mały pies – duży pies – duży pies

i naszą medianą jest kategoria „mały pies”, bo to ona znajduje się dokładnie pośrodku zbioru. Oczywiście w statystyce mamy na liczenie mediany specjalne wzory, a nie że biegamy po polu, zaczepiając wszystkie napotkane psy, i ustawiamy je w rzędzie w celu przeliczenia.

Średnia arytmetyczna

To trochę taka anglojęzyczna piosenka grana na koloniach w Mielnie – wszyscy ją znają, ale niewielu naprawdę rozumie. Idealnie sprawdzi się, gdy będziemy chcieli policzyć średnią długość naszych psów. Zmierzymy sobie, że jeden ma 45 centymetrów, drugi 23, a trzeci 22, dodamy wszystkie te centymetry, podzielimy przez liczbę psów (trzy) i wyjdzie nam, że w naszym zbiorze średnia długość psa to 30 centymetrów.

No dobra, ale tak właściwie po co nam trzy różne miary tego samego? Czy to jest tak, że ten rozkoszny zbiór miar tendencji centralnej możemy traktować jak cukiernię, to znaczy wybrać sobie z niego to, co lubimy najbardziej i co nam się najbardziej podoba? Otóż nie.

Wybór tendencji centralnej jest jak wizyta w cukierni, kiedy jest się na diecie – z wszystkich dostępnych opcji możemy wybrać tylko tę, na którą ktoś nam pozwoli. W tym wypadku to, na co pozwoli nam statystyka. Tak jest, to, jaką miarę tendencji centralnej wybierzemy, zależy od naszej skali pomiarowej (tutaj w tle powinien grać pełen uniesienia walc, jak w komediach romantycznych, kiedy to w ostatniej scenie wreszcie wszystko się ze sobą łączy, wszystkie zaginione szczeniaczki się odnajdują i okazuje się, że Karol jednak nie jest sukinsynem). Dla skal nominalnych możemy sobie policzyć wyłącznie dominantę, bo trzaskanie mediany i średniej byłoby zwyczajnie nieprawidłowe. Skale porządkowe polubią się z dominantą lub medianą, a najbardziej dopieszczają nas i tak wszelkie skale ilościowe, bo tam to hulaj dusza, piekła nie ma, możemy dowolnie wybierać, możemy sobie policzyć średnią i jeszcze to okrasić czułą kruszonką z mediany i posypką z dominanty.

Ale rozumiecie już, co to oznacza dla naszych nóg i łap? Skoro mierzymy te nogi i łapy na skali porządkowej, to śmiało możemy sobie poszukać dominanty i powiedzieć, że najczęściej w naszym zbiorze danych występują cztery nogi, niemniej liczenie tu średniej ma tyle samo sensu co liczenie średniej z koloru oczu lub rasy psa. No chyba że ktoś to naprawdę policzył i jest przekonany, że ma w domu 1,467 labradora, choć w takim wypadku musi być przygotowany na jeszcze jedno kluczowe pytanie:

Dzień dobry, czy pan upadł na głowę?

Słuchajcie, statystyka tak nie działa. Jasne, statystyka opisuje i przewiduje⁴⁰, ale warto pamiętać, że nie wnioskujemy statystycznie w sytuacji, gdy mamy dostęp do wszystkich elementów badanej populacji. To znaczy, gdy wychodzicie z psem na spacer, to możecie bez problemu sprawdzić, że pies ma – jedną, dwie, trzy, cztery – cztery łapy, a wy macie – raz, dwa – dwie nogi, zatem nie ma konieczności stosowania wyższych obliczeń statystycznych. Używanie wyższej statystyki w takim wypadku to tak, jakby z płonąca zapałką rozprawić się gaśnicą i z pomocą trzech zastępów straży pożarnej. *Hold your horses*, jak by to powiedział Irlandczyk.



Gdy zabraknie nam krytycznego myślenia, to łatwo dojść do oszałamiających wniosków, jak ten o liczbie nóg czy o tym, że kiedy wsadzimy głowę do piekarnika, a nogi do lodówki, to będzie nam w sam raz⁴¹. Zawsze tłumaczę moim studentom, że gdy cokolwiek liczą, to muszą

pomyśleć o dwóch rzeczach: po pierwsze, czy dany wynik ma sens, po drugie, co im ta informacja daje. Szalenie ważne jest to, by o tych naszych danych i wnioskach pomyśleć, podejść do nich z rozsądkiem. Chociażby po to, żeby nie zrobić z siebie idioty na rodzinnym grillu, kiedy to wywnioskujemy sobie, że jeśli mój pies ma patyk, a ja mam wołowinę, to średnio mamy po szaszłyku.

Zresztą nasza rola nie kończy się jedynie na myśleniu o tych danych czy też wsadzaniu ich w piękne tabelki. W końcu nadchodzi ten moment, za którym tęsknią absolutnie wszyscy, którzy na lekcji informatyki byli absolutnymi prymusami w tworzeniu logo szkolnej stołówki w WordArcie – wizualizacja danych. A jak wizualizacja danych, to także odpowiedź na pytanie, które spędza sen z powiek wielu demografom: czy 117% Polaków może się mylić?

Czy 117% Polaków może się mylić?

Najogólniej rzecz ujmując, to ja zupełnie się nie zgadzam z powszechnie panującą opinią, jakoby człowiek-Polak był z definicji chamem, co to je z kolanami na stole, a w autobusach podsiada staruszki w ciąży. Mnie tam Polska jawi się raczej jako kraj wyścielony pluszem, gdzie mieszka się w chatkach z lukru, a człowiek jest człowiekowi misiem koala, zawsze gotowym podzielić się swoim eukaliptusem. No bo przyjrzyjmy się przez chwilę tym cichym bohaterom polskiej codzienności, którzy drobiazgowo dbają o to, byśmy każdy dzień mogli witać z uśmiechem na naszych polskich psycholkach.

Weźmy na przykład tych ludzi Polaków, którzy z niezwykłą determinacją wspierają ludzi bez okien i rozpoczynają każdy dzień od poinformowania na Fejsie, jaka jest pogoda. I tych, którzy zawsze z najwyższą kulturą osobistą witają w autobusie wsiadających pasażerów tuż przy drzwiach i nie ma strachu, oni nie ruszą się z tego miejsca przez kolejne 13 przystanków, żeby przypadkiem nikt nie poczuł się pominięty. No i nie zapominajmy o ludziach-plastrach, którzy są czułym opatrunkiem na wszystkie rany odniesione w walce z codziennością i którzy świadczą swoje heroiczne usługi głównie w sklepach mięsnych, kiedy to pozwalają ludziom stojącym w kolejce za sobą podjąć świadomą decyzję i przed dokonaniem zakupu zawsze upewniają się, czy ta szyneczka jest świeża. A ta obok? A ile kosztuje? A ta pod nią? A ta za nią, obok tamtej? A czy smaczna? A ta tam z tyłu? A to dziękuję, przyjdę później.

Tylko przez narodową skromność nie wspominam o moim ulubionym temacie – człowieku-Polaku w samolocie, czyli o wisience na słodkim torcie dobrego wychowania, o listku mięty w mohito savoir-vivre'u. Taki człowiek-Polak ustawi się w kolejce odpowiednio wcześniej, żeby wskazać wam drogę do bramki, na waszym miejscu usiądzie, żeby wam je odpowiednio wygrzać, a czasem, w ferworze walki, da wam swoim bagażem po twarzy, żebyście przypadkiem na tej pędzącej autostradzie codzienności nie zapomnieli podumać chwilę nad okrucieństwem świata – dziś dostaniecie po mordzie od współpasażera, jutro od życia.

Tym bardziej krzywdzące wydają mi się opinie, że nie jesteśmy narodem gościnnym. Niech ktoś mi znajdzie człowieka-Polaka, który nie ma w szafie zrobionych ze sztucznego barana, kupionych w Zakopanem w 89 kapci dla

gości! Niemniej wiecie, jak jest – kaptcie z barana swoje, a statystyka swoje. Z wykresu wyświetlonego swego czasu w TVP Info jasno wynika, że 117% Polaków jest przeciwko przyjęciu uchodźców z państw Bliskiego Wschodu i Afryki.

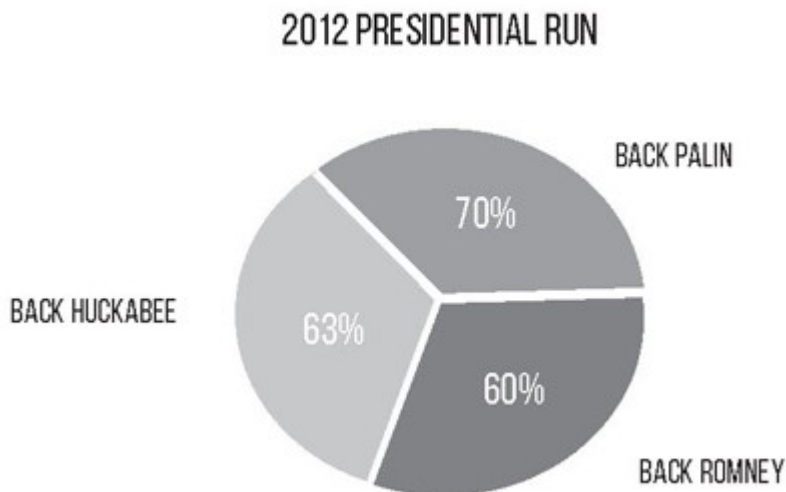


Autentyczny wykres, który pojawił się na antenie TVP Info 13.07.2017

A łącznie to w tej sondzie wypowiedziało się 143% naszego narodu, więc jeśli jesteście tymi ludźmi, którzy w czasie lekcji matematyki byli zajęci zlizywaniem lukru z wszystkich pączków w pobliskiej cukierni, to już tłumaczę: 143% to całkiem dużo⁴². I oczywiście wiem, co teraz sobie myślicie. Myślicie sobie, że to się przecież całkiem klawo składa, że nas jest tak dużo, bo teraz możemy sobie wszyscy pojechać do wesołego miasteczka i jak staniemy w kolejce do tej takiej karuzeli z łabędziami, to nikt nas nie wyprzedzi, nikt! No chyba że Irlandczycy.

Bo musicie wiedzieć, że ja kiedyś moim miłym studentom z kraju koniczyny pokazałam tabelę krzyżową, a to taka tabela, w której jest tyle liczb naraz, że aż człowiek nie wie, gdzie patrzeć, a jakby tego koszmaru było mało, to jest tam też tekst, co oznacza, że człowiek musi jednocześnie i czytać, i myśleć, czyli trochę zgroza. A więc pokazałam im tę tabelę, a podekscytowana przy tym byłam jak labrador w fabryce kiełbasy. Zadanie było następujące: proszę sobie przeczytać tabelę i potem na jej podstawie powiedzieć mi, ile procent ludzi w Irlandii ma bardzo dobre zdrowie. Oni się zastanowili, policzyli i oświadczyli mi, że 370,8% ludzi

w Irlandii ma bardzo dobre zdrowie, i w tamtym momencie bardzo wyraźnie poczułam, że ja do tych 370% ludzi na pewno nie należę. I wiecie, to nie tylko dlatego, że zrozumiałam, że oto w tej kolejce do karuzeli z łabędziami to na bank nas wyprzedzą ci pełni wigoru Irlandczycy. Właściwie to jest jeszcze gorzej, bo Irlandczykom po ich szalenie zdrowych piętach deptać będzie 193% Amerykanów, którzy głosowali w wyborach prezydenckich w 2012 roku.



Wykres pojawił się na antenie Fox, 20.11.2009

No i widzicie, jak to jest, już się człowiek podjarał, że wraz ze swoimi 143% narodu ma jakiejkolwiek szanse w kolejce do karuzeli, a tu okazuje się, że w porównaniu z innymi nacjami to my jak zwykle mamy pod górkę, i to w sandałach z rzemyków i za ciasnym berecie. Niemniej mnie osobiście trochę pociesza to, że nie tylko w naszym kraju znalazł się taki geniusz, który uwierzył, że z procentami na wykresach jest jak z sosem do klusek – im więcej, tym lepiej. A wiecie, co by mnie pocieszyło jeszcze bardziej? Kluski z sosem. Ewentualnie ten moment, kiedy przestano by oszukiwać nas za pomocą wizualizacji danych. Poprawcie berety, nauczę was, jak prawidłowo interpretować wykresy.

Tu jest jak w monopolowym, tu wszystko jest pyszne!

Wyobraźmy sobie taką sytuację: sobota, dzień tak przyjemny, że życie pluszem wykłada ziemię, po której stąpacie, miodem dosładza gorzką pigułkę dni roboczych, z czułością nosi was po świecie tak dobrym

i czystym, że słycać w nim tylko śpiew ptaków, śmiech dziecka i gdzieś tam daleko rozkoszny szelest spadającego kursu franka. W tej rozkosznej atmosferze otwieracie sobie jakiś plik z danymi, wybieracie dostępne opcje wizualizacji i widzicie: wykres słupkowy, kołowy, liniowy, pudełkowy, histogram... I wiem, co sobie teraz wszyscy myślicie. Myślicie sobie: o mój Boże, tu jest jak w monopolowym, tu wszystko jest pyszne!

Założmy jednak, że nie macie czasu spróbować wszystkiego z monopolowego, no nie jesteśmy już na studiach. W jaki sposób wybrać najwspanialszy wykres z tej życiowej cukierni wizualizacji danych? No musimy zacząć od pytania, jakim rodzajem skali pomiarowej jest przedstawiona zmienna na owym wykresie. Ewentualnie, jeśli jesteśmy dopiero na samym początku naszej drogi statystycznego świra, musimy zacząć od pytania, co to w ogóle jest skala pomiarowa. I zmienna. I po co to komu. I za jakie grzechy. I jeszcze od wielu innych pytań, które często zadają mi studenci⁴³.

Zacznijmy od tej zmiennej. Gdyby zmienna była psem, to najpewniej byłaby labradorem, albowiem to jest taki koncept o bardzo prostej definicji i bardzo trudnej nazwie.



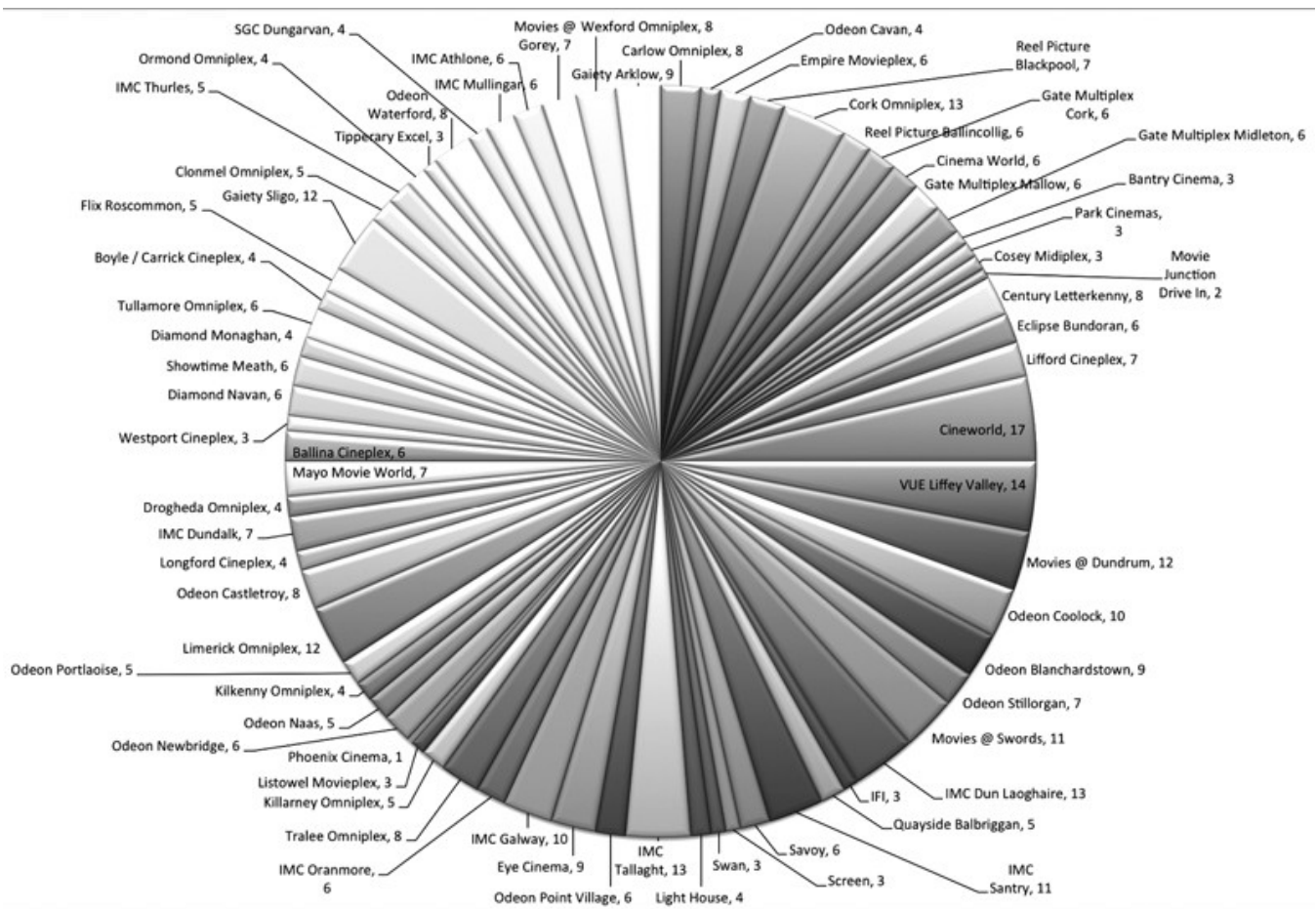
Zmienna to właściwość lub cecha, pod względem której elementy populacji różnią się od siebie. Zmienną jest więc na przykład płeć (przyjmuje dwie wartości – kobieta lub mężczyzna), wiek (przyjmuje dowolne wartości liczbowe – np. 5, 80 czy 95 lat), wyznawana religia, zawód, wzrost, inteligencja i wiele innych. Zmienna nie musi opisywać ludzi. Są to również wszelkie właściwości instytucji (na przykład ich rodzaj – szkoła, urząd, sklep; typ – publiczna czy prywatna; wielkość; kolor elewacji czy liczba pięter), zwierząt (gatunek, umaszczenie, stopień

udomowienia czy nawet imię) i pozostałych elementów otaczającego nas świata.

Zmienne mogą być mierzone na różnych skalach pomiarowych, o których mówiliśmy w poprzednim rozdziale – na skali nominalnej, porządkowej, przedziałowej lub ilorazowej. Zarówno dobór miar statystycznych, jak i sposób prezentacji danych zależą od typu zmiennej, którą badamy. Tak jest, od rodzaju skali zależy sposób analizy, ale również to, jakiego rodzaju wykresu użyjemy do jej opisu. Zmienne ilościowe, czyli przedziałowa i ilorazowa, polubią się z histogramami liczości i gęstości, wykresami pudełkowymi lub wykresami liniowymi w przypadku trendów czasowych. Skale jakościowe, czyli nominalna i porządkowa, są bardziej grymaśne – w ich przypadku mamy do wyboru głównie wykresy słupkowe i kołowe. Z czego ten ostatni to taki Fiat Multipla świata statystyki – niby ładnie wygląda, ale nikt go nie szanuje.

O tym, że wykres kołowy to Fiat Multipla świata statystyki

Podstawowym celem jakiegokolwiek wizualizacji danych jest ułatwienie naszym odbiorcom życia tak żeby mogli sobie rachuciachu spojrzeć na wykres i w mig pojąć, co tam się dzieje. Wykres kołowy wydaje się idealny, gdy chcemy pokazać panu prezesowi, że udziały w rynku naszej firmy są znacznie większe niż całej konkurencji razem wziętej. Pan prezes łatwo zsumuje wszystko do 100%, łatwo zobaczy, komu w tym zestawieniu wiedzie się lepiej i że z pewnością jemu. No dobrze, ale co w przypadku, gdy wykres kołowy przedstawia dane, które mają, no nie wiem, trochę więcej kategorii?



Boże mój, człowiek zaczyna analizować ten wykres z niemowlakiem na rękach, a gdy kończy, to ten ma pretensje, że nie przyszło się na jego osiemnastkę. Trudno się w tym wszystkim połapać, no nie? A tak naprawdę nawet gdy ograniczymy liczbę kategorii, kłopot nie zniknie. Udowodnili to William S. Cleveland i Robert McGill (1985).

Tych dwóch miłych naukowców przeprowadziło szereg eksperymentów, które miały sprawdzić, w jaki sposób radzimy sobie z przetwarzaniem wykresów. Badanym pokazywano dwie figury geometryczne o różnej wielkości, a następnie mieli oni oszacować proporcje między tymi figurami (w procentach, czyli o ile jedna z nich jest mniejsza od drugiej). Wyniki pokazały, że zdecydowanie najłatwiej nam szacować proporcje i zależności między kategoriami, gdy są one pokazane wzdłuż jednej i tej samej skali. W rankingu najprostszyc zadań nasze umiejętności szacowania różnic pomiędzy kątami lub powierzchniami (co robimy w wypadku wykresów kołowych) były dopiero na czwartym i piątym miejscu (z siedmiu)⁴⁴. Innymi słowy, łatwiej radzimy sobie z porównywaniem skal długości niż powierzchni pól. Wie o tym każdy, kto próbował kiedyś zdecydować, który

z dwóch ostatnich kawałków pizzy jest największy, by ostatecznie wezwać na pomoc geodetę.

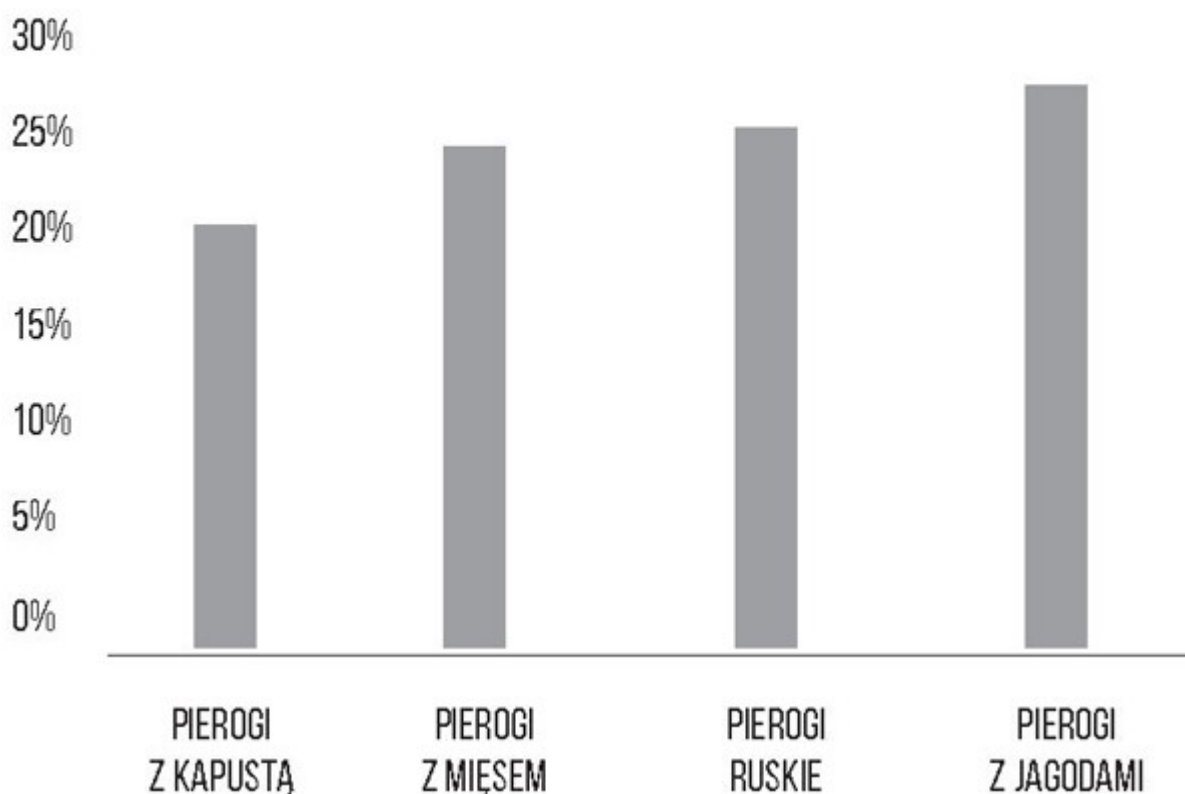
Słuchajcie, to nie jest tak, że chciałabym propagować politykę nienawiści względem wykresów kołowych. Musicie wiedzieć, że eksperymenty Clevelanda i McGilla nie były jedynymi w tym temacie, a pozostałe badania, które mierzyły efektywność różnych metod wizualizacji, okazały się znacznie mniej surowe względem wykresów tego typu, zwłaszcza w przypadku prezentowania znacząco różniących się od siebie wartości (Spence i Lewandowski 1991, Simkin i Hastie 1987, Hollands i Spence 1998). Niemniej zwracam uwagę na to, że bardzo często przedstawienie tych samych danych na wykresie słupkowym znacznie ułatwia odbiorcy życie, jeśli chodzi o szybkość i trafność uszeregowania określonych kategorii. O, sami spróbujcie przeprowadzić taki mikroeksperyment. Poniższy wykres przedstawia częstotliwość występowania określonych rodzajów pierogów na niedzielnym obiedzie. Celowo usunęłam z niego etykiety wartości. Spróbujcie uszeregować te kategorie od największej do najmniejszej:

GĘSTOTLIWOŚĆ WYSTĘPOWANIA OKREŚLONYCH
PIEROGÓW NA NIEDZIELNYM OBIEDZIE



A teraz pokażę wam te same dane na wykresie słupkowym i poproszę o to samo – o uszeregowanie kategorii w kierunku od najliczniejszej.

CZĘSTOTLIWOŚĆ WYSTĘPOWANIA OKREŚLONYCH PIEROGÓW NA NIEDZIELNYM OBIEDZIE (W %)



I co, łatwiej? Szybciej? A pomyślcie tylko, co możecie zrobić z tymi kilkoma sekundami, które zaoszczędziliście przy odczytywaniu danych z wykresu słupkowego zamiast kołowego! Książkę możecie napisać, tańczyć w deszczu, kochać tak, jak jeszcze nigdy nikt...

I w gruncie rzeczy to ma znaczenie – zarówno taniec w deszczu, jak i to, jak szybko odbiorcy poradzą sobie z przetwarzaniem informacji, które im pokażemy na wykresie. Tak naprawdę jedną z największych zalet wykresu kołowego jest to, że nazywa się go również tortowym, a torty są jednak pycha. Wykres kołowy nie jest więc najskuteczniejszą formą prezentowania danych, niemniej jeśli ktoś z jakiegoś powodu jest do niego szalenie przywiązany, to podaję kilka sposobów na to, by był on trochę przyjemniejszy w odbiorze. Po pierwsze, posegreguj kategorie rosnąco. Po drugie, spraw, by największa kategoria zaczynała się na godzinie 12. I po trzecie, umieść etykiety z wartościami blisko wycinków, aby nikt nie musiał błędzić wzrokiem po całym wykresie, żeby zrozumieć, o co chodzi. Po zastosowaniu tych zasad wciąż takiemu wykresowi będzie – wbrew

szalenie obiecującej nazwie – daleko do pięknego tortu czekoladowego, ale też nie będzie to już brownie z fasoli z kruszonką ze szparagów.

Porady dotyczące tego, w jaki sposób sprawić, by inni ludzie mniej cierpieli przy odczytywaniu naszych danych, dotyczą zresztą nie tylko rodzaju wykresu, lecz także wszelkich innych jego charakterystyk: doboru kolorów, trójwymiarowości czy liczby kategorii i zmiennych pokazanych na jednym obrazku.

Co o twoim wykresie pomyśli cocker-spaniel?

Bo to, o czym wielu grafików i analityków zapomina, to to, że wszelkie wizualizacje powinny ułatwiać odbiorcy zrozumienie i interpretację danych – tak aby mógł to zrobić jak najszybciej i jak najtrafniej. Dobry wykres nie musi stanowić katalogu wszystkich naszych ulubionych kolorów ani zaspokajać naszej potrzeby malowania wykresowych słupków na kolor dziko rosnących po ulewnych deszczach filodendronów, ewentualnie transparentnej musztardy o zachodzie słońca. Wykres rysujemy po to, by ludziom było łatwo zrozumieć, co się dzieje w danych, a nie po to, by pochwalić się, że nie jesteśmy kornikiem i widzimy w trójwymiarze. Właściwie jedynym zwierzęciem, o którym powinniśmy myśleć w trakcie wizualizacji danych, jest cocker-spaniel.

Serio, kiedy już wybierzemy wszystkie kolory, odcienie i inne dodatki do naszego wykresu, zawsze warto zadać sobie na końcu to jedno kluczowe pytanie: a co by było, gdyby ten wykres oglądał cocker-spaniel? Albo ktoś, kto nie widzi kolorów, kto ma problem z przetwarzaniem głębokości obrazu, ewentualnie po prostu siedzi w ostatnim rzędzie na konferencji (bo najbliżej do ciasteczek) i jeszcze mu słońce świeci w ekran?

Zróbmy wszystko, by obcowanie z naszym wykresem było dla innych ludzi tak pluszowe i bezproblemowe, jak tylko to możliwe. Oto kilka reguł, których warto przestrzegać, by nikt nie dzwonił na policję (albo co gorsza – na pogotowie) na widok waszych wizualizacji:

Zasady wizualizacji danych:

1. Wykres to nie walentynka – musi być **PODPISANY**. Dotyczy to zarówno tytułu, jak i opisanie odpowiednich osi.

2. Wszystkie elementy wykresu (np. kolory, kształty punktów, tekstury) są jak dzieci – muszą być odpowiednio NAZWANE. Odpowiednio to znaczy tak, by nikt nie miał żadnych wątpliwości, czego dotyczą i co oznaczają. Żeby potem nie było tak jak z niektórymi ludźmi, co to nie mają żadnych wątpliwości, że powinni się nazywać Zarządca Niewieścich Serc i Władca Wzruszeń, a matka przy chrzcie zdecydowała, że Jędrzej.

3. Pamiętajcie, że 25% dużej pizzy to nie to samo co 25% małej pizzy – zawsze podawajcie LICZBĘ WSZYSTKICH OBSERWACJI N, które obejmuje wykres.

4. Wizualizacja to nie aktywność fizyczna – jej celem nie jest utrudnienie ludziom życia. Uprośćcie swoim odbiorcom zrozumienie wykresu – na przykład dzięki dodaniu ETYKIET Z WARTOŚCIAMI.

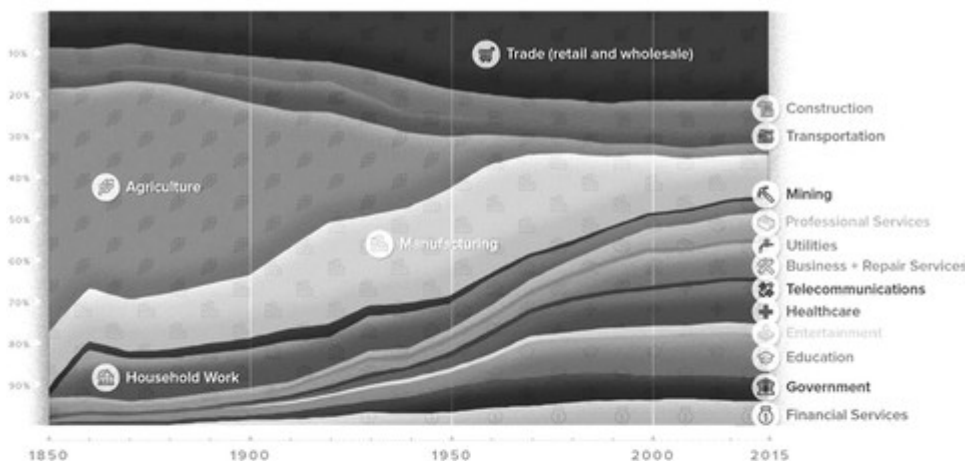
5. Wykres to nie pisanka – KOLORY stosujcie ostrożnie.

6. Myślcie z czułością o ludziach, którzy stają w płomieniach na widok liczb – pokazujcie na wykresie dokładnie tyle, ile jest niezbędne, i **PODKREŚLAJCIE** to, co najważniejsze. Wykres to nie studniówka – nie potrzebuje brokatu, udziwnień i innych fiu-bździu.

Pamiętajcie również, że tego wykresowego kwiatu jest pół świata. Oprócz znanych nam wszystkim wykresów kołowych, słupkowych czy radarowych mamy jeszcze mnóstwo innych opcji, takich jak wykres bąbelkowy czy pudełkowy, przez niektórych zwany uroczo wykresem pudełkowym z wąsami. To, jaki typ wykresu wybierzemy, zależy od rodzaju naszej skali pomiarowej, formy publikacji (w raportach możemy sobie pozwolić na trochę więcej niż w prezentacjach, gdzie nasi odbiorcy mają tylko kilka sekund, by szczić, co przedstawia slajd⁴⁵), jak również – od naszej wyobraźni. Tam, gdzie jest to uzasadnione i nie zakłóca odbioru, warto zawierzyć Davidowi McCandlessowi, który stwierdził, że „informacja jest piękna”⁴⁶. Popatrzcie na przykład na to wizualizacyjne dzieło sztuki, które przedstawia zmiany w sektorach amerykańskiej gospodarki na przestrzeni 150 lat:

VISUALIZING 150 YEARS OF U.S. EMPLOYMENT HISTORY

How sector shares of jobs have changed over time



A Changing Workforce

Over time, the job landscape tends to shift dramatically for almost any economy. In the above graph, we show the share of U.S. employment provided by each sector between the years of 1850 and 2015.

How will AI and automation shape the future of employment?



Agriculture's share of U.S. employment was close to 60% in 1850, but today it represents just 3% of jobs.

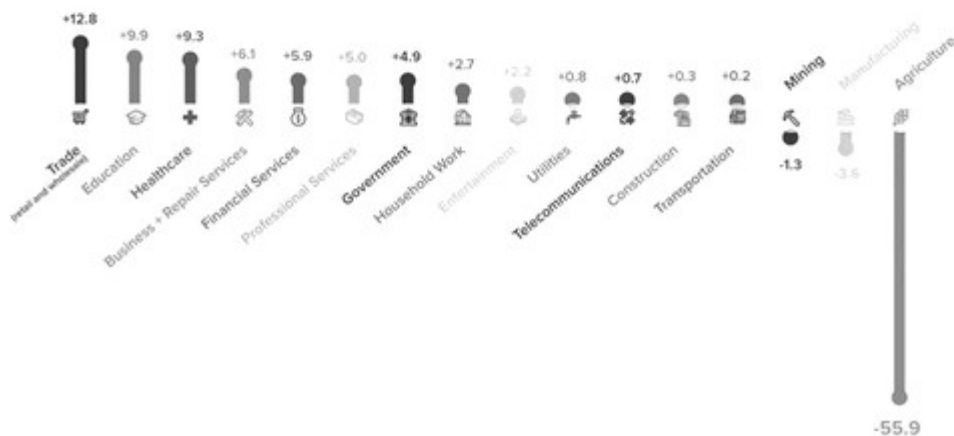
While these declines seem significant and fast, in recent years China has seen an even more stunning change: between 1990-2015, a full 1/3 of China's workforce moved out of the agriculture sector.



Manufacturing represented 26% of U.S. employment in 1960, but it is below 10% today.

EMPLOYMENT SHARE CHANGE 1850-2015

PERCENTAGE POINTS



Tutaj wszystko się zgadza: są jasno opisane kategorie i osie, ładnie widać zmiany, które zaszły na przestrzeni czasu, a i całość jest miła i atrakcyjna wizualnie⁴⁷. Ach, gdyby ten wykres był ciastem, to byłby najdoskonalej wyrośniętą babką drożdżową polaną słodkim lukrem. Niestety, za szybą stoją zwykle zakalce.

Zakalce świata wizualizacji danych

Marzy mi się kiedyś taka sytuacja, że biega sobie reporter po ulicy i pyta przechodzących ludzi o najbardziej podziwianych przez nich Polki i Polaków, i na tej jego liście trwa jakaś najdziwniejsza impreza świata – Zenek Martyniuk miesza się z Janem Pawłem II, Violetta Villas tańczy z Józefem Piłsudskim, no i oczywiście zawsze znajdzie się ten jeden człowiek, który będzie twierdził, że na liście powinien się znaleźć Johnny Depp, bo ten to na bank jest w 1/18 nasz, ojczysty. Dajcie spokój, mamę oszukasz, tatę oszukasz, ale naszych narodowych genów nie oszukasz. Przecież on totalnie wygląda jak sprzedawca precli z Kędzierzyna-Koźła, jest tylko kwestią czasu, aż będzie śpiewał *Barkę* w naszej narodowej telewizji i opowiadał, że w sumie to jest w 1/3 pierogiem. Niemniej wiecie, kto zdecydowanie powinien się znaleźć na liście, zaraz po Janie Pawle II i człowieku-preclu? Dr hab. inż. Przemysław Biecek. Nasz polski David McCandless, tylko że na odwrót.

Założył on fundację Smarter Poland, która zajmuje się propagowaniem wiedzy o matematyce i statystyce, a po godzinach tropi najgorszych przestępców statystycznego świata – nikczemników, którzy nadużywają liczb, i łachudrów, którzy oszukują za pomocą wykresów. Jedną z najwspanialszych inicjatyw fundacji jest coroczny konkurs na Najgorszy Wykres Roku. To w sumie takie Złote Maliny statystyki, tylko że w internecie, dzięki czemu nikt z nas nie musi wychodzić z piwnicy do tego niebezpiecznego świata (zwanego też prawdziwym), w którym nikt nie śmieje się z wyborczego żartu o tym, że wchodzi całka oznaczona do pociągu, a to nie jej przedział.

Polecam wam zajrzeć na bloga Smarter Poland, by zobaczyć najbardziej spektakularne zakalce świata wizualizacji, a gdyby komuś bardzo zależało na wygraniu tego konkursu, to mam dla was kilka prostych sposobów, jak narysować wykres oszusta w taki sposób, by mieć spore szanse na zwycięstwo. Ewentualnie możecie tę część rozdziału potraktować jako krótki poradnik, jak zabić Janinę.

1. Wersja dla leniwych: wykres słupkowy, którego skala Y nie zaczyna się od 0

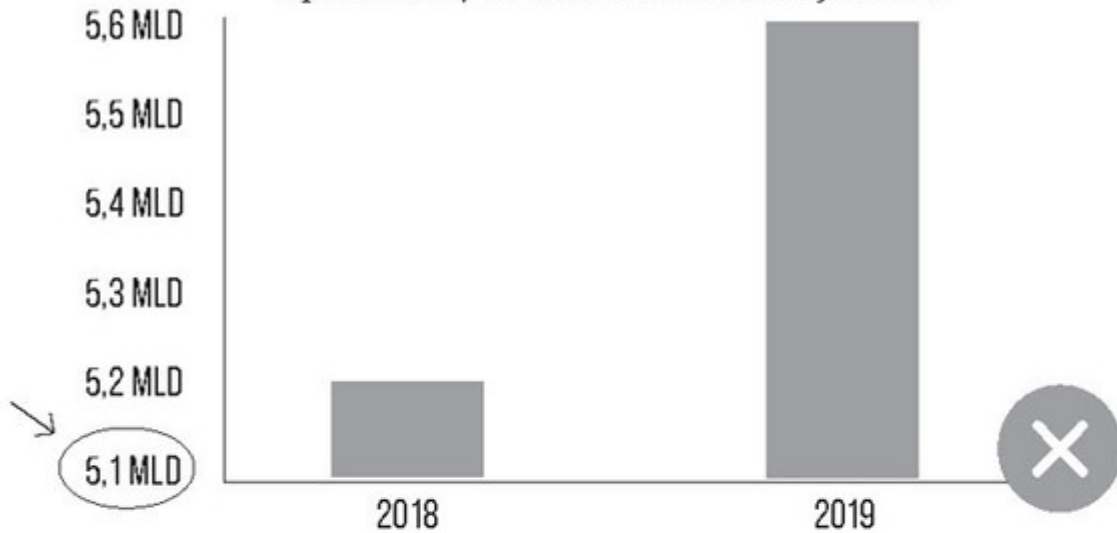
Obawiam się, że ludzie, którzy oszukują wykresami w ten właśnie sposób, ukończyli razem z wszystkimi labradorami szkołę Najgorszych Przestępców Świata. Jest to wersja dla najbardziej leniwych oszustów i polega na tym,

że rysujemy wykres słupkowy, a potem przesuwamy trochę jego skalę Y, tak by nie zaczynała się od 0, a tym samym by różnice między kategoriami wyrosły przepięknie niczym racuchy w ciepłe. Przykład? A proszę bardzo.

Ministerstwo Zdrowia postanowiło swego czasu pokazać oszałamiający wzrost przewidywanych nakładów finansowych na łódzki NFZ. Niestety nie mieli wystarczająco dużo pieniędzy w skarbcu, by zaoferować szokujące podwyżki, dlatego musieli się zadowolić skalą Y, która zaczynała się od 5,1 miliarda złotych. Dobrze to wymyślili, bo tym samym różnica pomiędzy dwoma słupkami, która wynosiła ledwie 0,4, na wykresie jawiła się jako gigantyczna (wykresy na następnej stronie). Przynajmniej do momentu, aż jakaś psuja dobrego samopoczucia nie postanowiła narysować tego wykresu poprawnie, co zniweczyło cały spektakularny efekt osiągnięty przez naszych geniuszy wizualizacji. Państwo z Ministerstwa Zdrowia, wasz wykres był bardzo chory, ale nie martwcie się, uleczyłam go. W tym wypadku leczenie nie było zbyt skomplikowane, wymagało tylko odpowiedniej skali i odrobiny przyzwoitości.

TE SAME DANE, DWA RÓŻNE WYKRESY

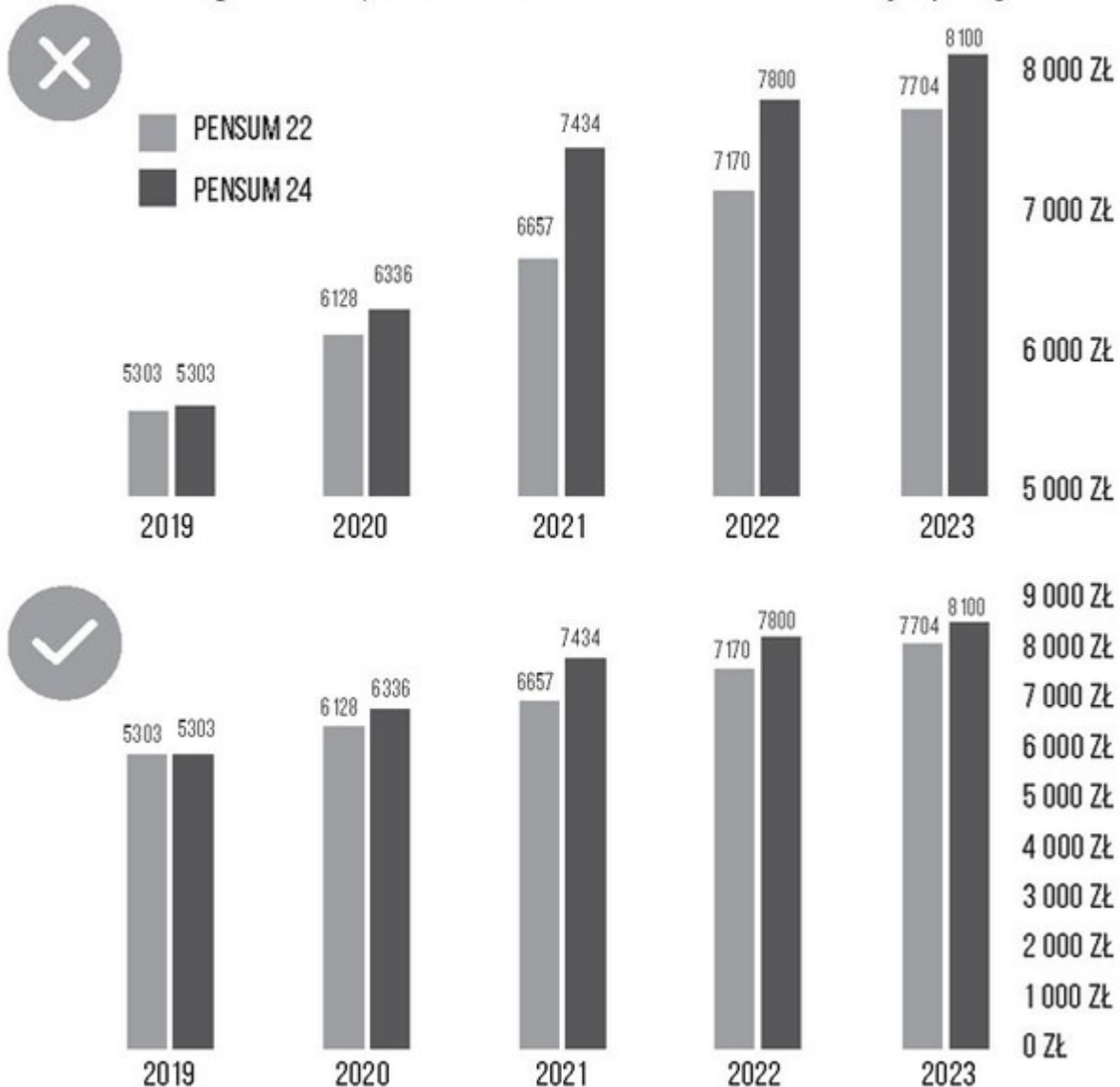
PLAN FINANSOWY ŁÓW NFZ (ŁÓDZKI)
Opublikowany na Twitterze NFZ dnia 15.10.2018



Teraz na pewno zastanawiacie się, od kogo Ministerstwo Zdrowia nauczyło się rysowania tak złych wykresów. Już odpowiadam: najpewniej od kolegów z Kancelarii Premiera, którzy swego czasu postanowili załatać dziury w pensjach nauczycieli, używając podobnego narzędzia, tj. zabawy skalą.

TE SAME DANE, DWA RÓŻNE WYKRESY

ŚREDNIE MIESIĘCZNE WYNAGRODZENIE NAUCZYCIELA DYPLOMOWANEGO I JEGO STOPNIOWY WZROST WRAZ ZE WZROSTEM PENSUM ORAZ PODWYŻKAMI BUDŻETOWYMI
Opublikowany na Twitterze Kancelarii Premiera dnia 05.04.2019



Taka forma wizualnej manipulacji, kiedy to istnieje niezgodność między wielkościami przedstawionymi na wykresie a rzeczywistą różnicą wynikającą z danych, nazywa się **WSPÓŁCZYNNIKIEM KŁAMSTWA** (ang. *lie factor*). Pierwszy napisał o niej profesor Edward Tufte, prawdziwy miłośnik wizualizacji danych i tworca definicji rozmaitych współczynników⁴⁸. Wzór na współczynnik kłamstwa jest prosty i dość intuicyjny: jest to stosunek efektu widocznego na grafice do efektu

wskazywanego przez dane, na podstawie których powstała grafika. Dla przykładu powyższy wykres przedstawiający zarobki nauczycieli ma współczynnik kłamstwa przekraczający 350%, czyli różnice pokazane na wykresie są ponadtrzykrotnie większe, niż wynikałoby to z danych.

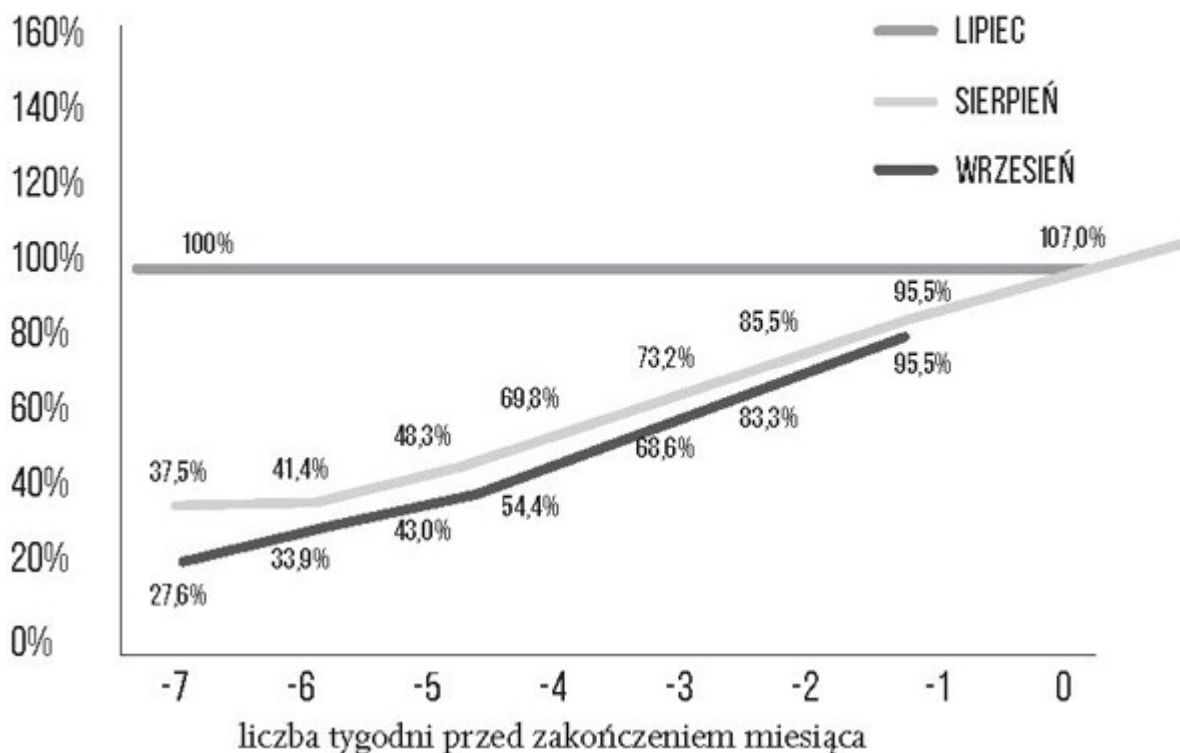
Niemniej nie ma strachu, jeśli ktoś ma ochotę na bardziej wyrafinowane techniki manipulacji, to nasz miłośnik tworzenia definicji, profesor Tufte, ma dla was jeszcze jedną propozycję – wykres śmieciowy (ang. *chartjunk*).

2. Wersja dla miłujących splendor: wykres śmieciowy

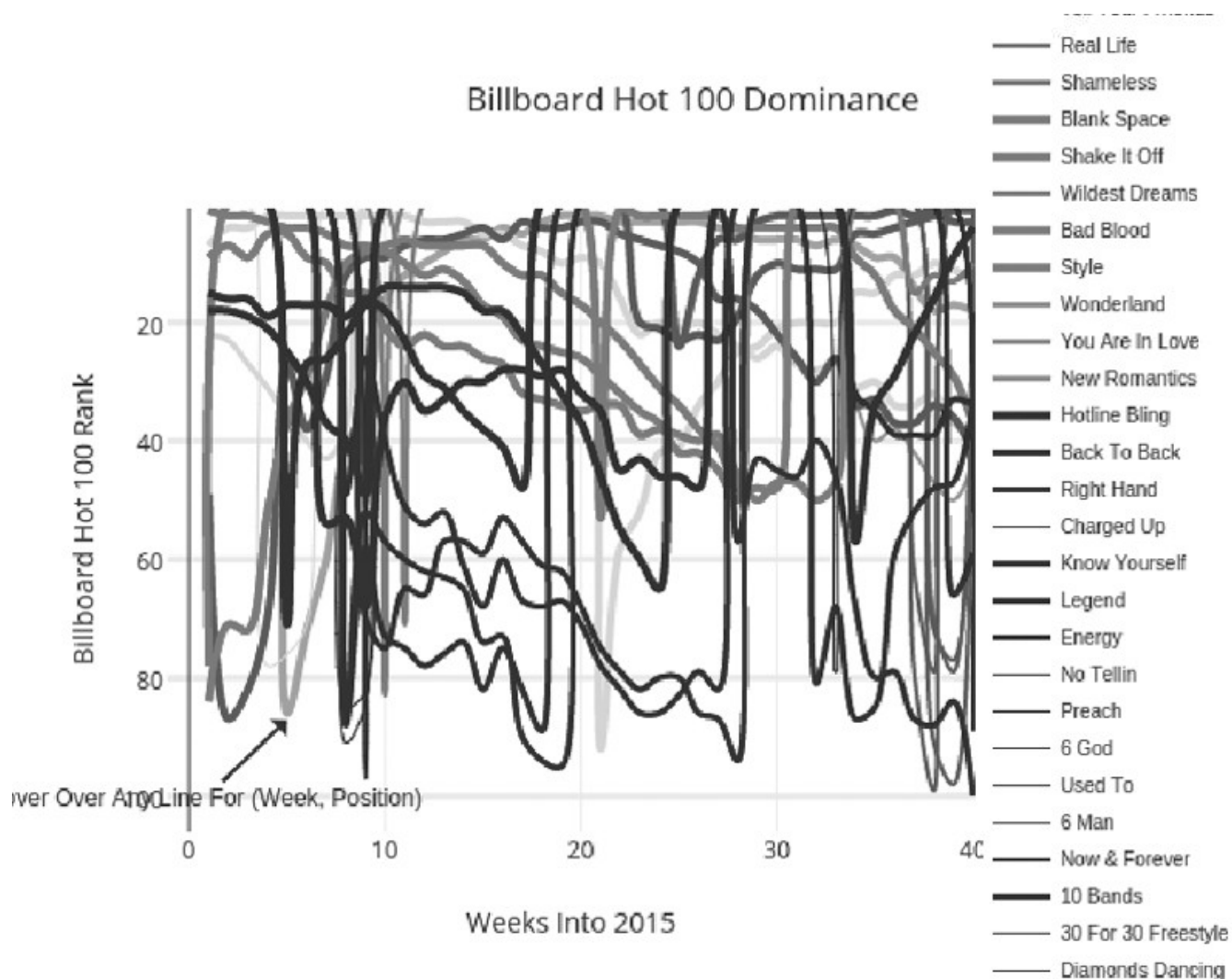
Powiem wam, że mam spory kłopot z tą nazwą „śmieciowy”. Odnoszę wrażenie, że ona zupełnie nie oddaje sprawiedliwości autorom owych wizualizacji, którzy są jak trzyletnie dzieci ubierające choinkę, że tutaj sobie bombki zawiesimy we wszystkich kolorach tęczy, tutaj trzaśniemy kilka błyszczących łańcuchów, watę imitującą śnieg i jeszcze postawimy sztuczną łąsicę pod drzewkiem, bo dlaczego nie. Innymi słowy, wykres śmieciowy to taki, w którym czytelny przekaz ginie pośród nadmiaru rozprasających elementów i uduwnień⁴⁹. Ach, patrzcie na te perełki!

Spójrzmy dla przykładu na kandydata numer 1. Gdyby żył w XIX wieku, to napisałby *Nad Niemnem*, niemniej w obecnej sytuacji musiał zadowolić się tworzeniem tytułów wykresów. Szczęśliwie nie ma strachu – to, czego nie zrozumiemy z opisu, bez problemu wyczytamy z danych i samego wykresu. OH, WAIT.

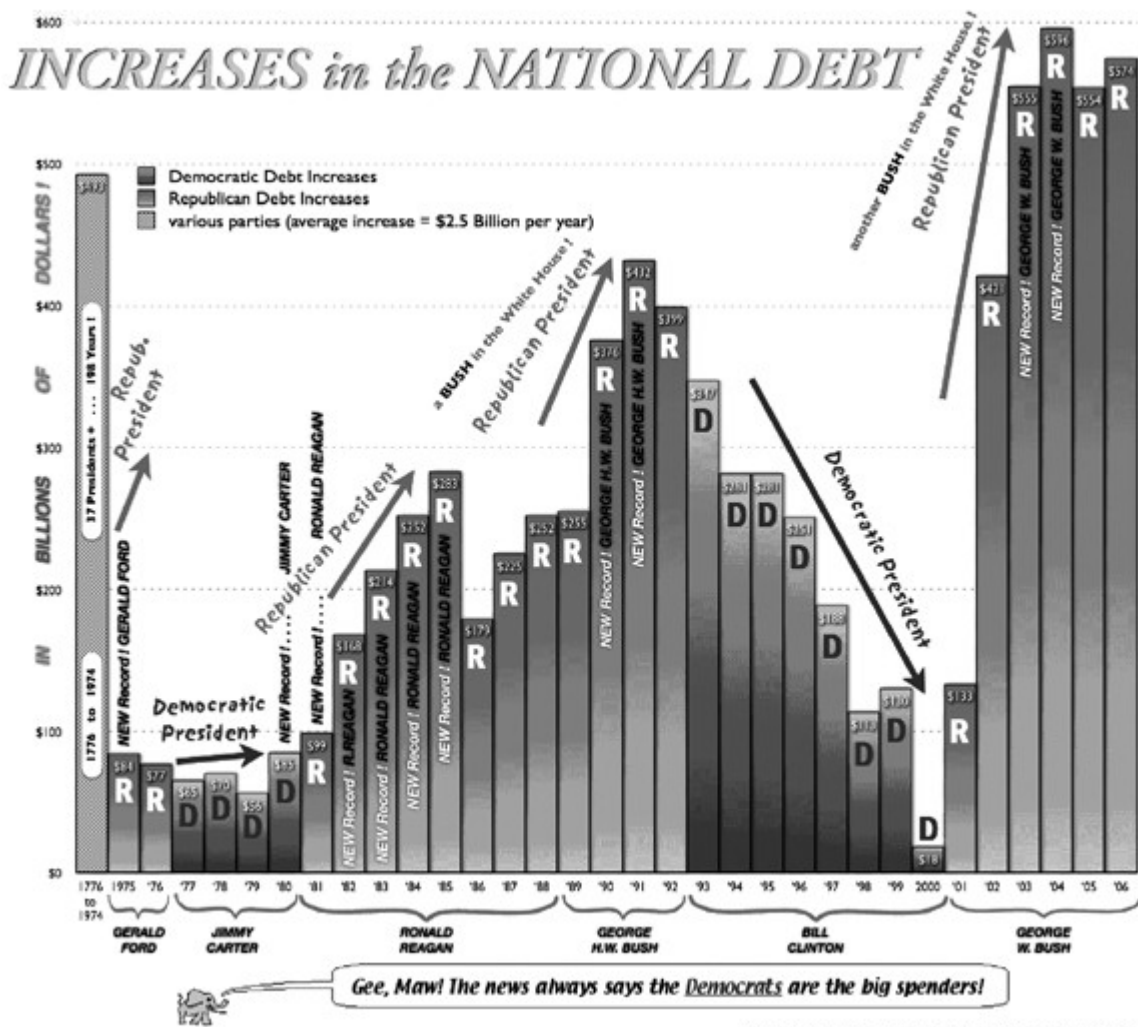
ZAAWANSOWANIE LICZBY SPRZEDANYCH WYCIECZEK W SIERPNIU I LIPCU (W %) NA TLE LICZBY SPRZEDANYCH W LIPCU (LINIA 100%) NA KONIEC POSZCZEGÓLNYCH TYGODNI PRZED ZAKOŃCZENIEM MIESIĄCA (Na podstawie danych Enovatis)



Kandydat numer 2 pozwoli wam zrozumieć, które utwory muzyczne były najpopularniejsze w 2015 roku. No chyba że nie.



Z ostatniego kandydata nie wolno nam się śmiać, bo mam w sobie głębokie przekonanie, że tak dramatyczna miłość do wszelkich linii to musi być jakaś choroba. Chciałam nawet policzyć, czy udałoby nam się znaleźć w amerykańskim budżecie jakąś kwotę na leczenie autora, ale wszystko wskazuje na to, że w chwili obecnej mają tam dość duży problem z długiem publicznym. To znaczy, chyba mają, ewentualnie nie mają, w sumie nie wiadomo.

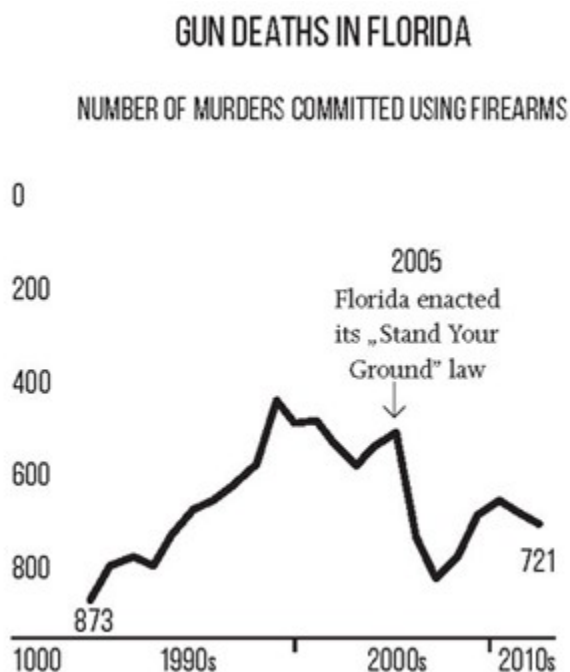


Czego nas uczą te wszystkie przykłady? Po pierwsze, że nie należy nigdy rysować wykresów po pijaku. Po drugie, że rysowanie wykresów to nie wizyta w cukierni – umiar jest wskazany. No chyba że należycie do trzeciej kategorii autorów – tych, którzy nie mają już absolutnie żadnej litości.

3. Wersja dla barbarzyńców: wykresy oszuści

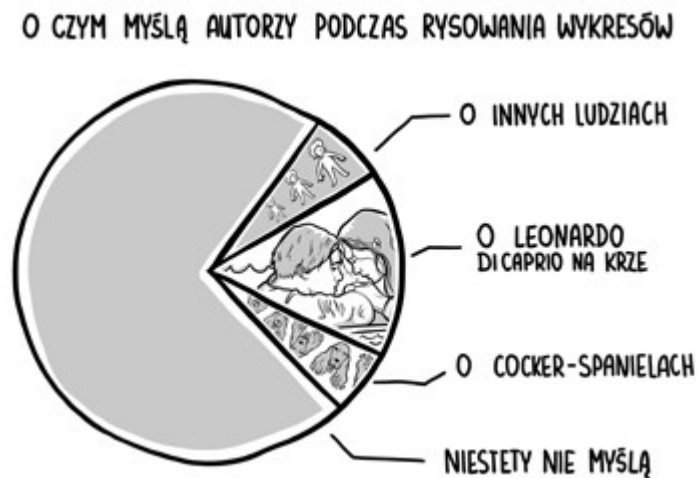
I znowu człowiek się podjarał, że jego własne ministerstwa i instytucje rządowe mają tak zaawansowane umiejętności w zakresie statystycznej chirurgii plastycznej, potrafią tak wspaniale podrasowywać wykresy, tu podciągnąć, tu podkolorować, by wszystko wyglądało idealnie, a potem przychodzi sobie takie Ministerstwo Sprawiedliwości z Florydy i wszystkich nas zawstydzają.

Historia była taka: w 2005 roku na Florydzie wprowadzono prawo *Stand your ground law*, które zezwalało właścicielowi na zastrzelenie każdego, kto wtargnął na jego posesję. Moglibyście więc pomyśleć, że to nie do końca korzystna sytuacja, niemniej chwilę później Agencja Reutersa opublikowała wykres, na którym pokazała, jak zmieniała się na Florydzie liczba śmiertelnych postrzałów na przestrzeni lat. Na pierwszy rzut oka wszystko wyglądało szalenie optymistycznie, albowiem w 2005 roku liczba postrzałów opadła drastycznie niczym sernik zbyt wcześnie wyciągnięty z piekarnika. Czy nie?



Jeśli przyjrzą się dobrze, to zobaczą, że skala jest do góry nogami, co sprawia wrażenie, jakoby liczba postrzałów zmalała, mimo iż tak naprawdę wzrosła o prawie 50% na przestrzeni kolejnych dwóch lat. Przyznam, że swego czasu spędziłam dość sporo czasu przy telefonie, starając się przekonać Ministerstwo Sprawiedliwości, że autorka tego wykresu, Christine Chan, zasługuje na najwyższy wymiar kary, to znaczy ja rozumiem, że dożywocie może być uznane za pewną przesadę, ale czy mogę chociaż zaproponować 25 lat słuchania największych przebojów Ich Troje w aranżacji country? Niemniej Christine Chan twierdzi, że jej intencją nie było wprowadzenie nikogo w błąd, tylko pokazanie rosnącej liczby przypadków śmierci za pomocą czerwonej (jak krew) powierzchni wykresu⁵⁰.

Niestety, nawet jeśli sam pomysł wydaje się niegłupi, to rozbił się o ostre krawędzie ludzkiej percepcji, po czym podryfował w dzikie rejony braku logiki. A o tym, że to nie jest miła sytuacja, wie absolutnie każdy, kto oglądał *Titanica*. Co w sumie całkiem nieźle podsumowuje wszelkie wytyczne dotyczące wizualizacji danych – podczas rysowania jakiegokolwiek wykresu myślcie z miłością o wszystkich ludziach, którzy o liczbach myślą nienachalnie entuzjastycznie. O cocker-spanielach. No i o Leonardo DiCaprio dryfującym na samotnej krze. Co całkiem nieźle przedstawia poniższy wykres:



Dobra, ale chwilowo mówiliśmy o mierzeniu rzeczy, które w jakiś sposób są dla nas intuicyjne – wynagrodzenie, procent osób lubiących coś tam, a nie lubiących czegoś innego, lub liczba nóg. W gruncie rzeczy nie ma w tym nic skomplikowanego – wchodzimy na konto i widzimy, ile zarabiamy, choć nasz kot wciąż nie wie, dlaczego niewystarczająco, by codziennie karmić go kawiozem z tuńczyka. Albo liczymy, że jedna noga i druga noga, i nawet nie potrzebujemy do tego liczydła, no chyba że jesteśmy chrząszczem. Ale co z trochę bardziej abstrakcyjnymi konceptami – na przykład jakością życia czy zadowoleniem z małżeństwa? Czy szczęście też możemy zmierzyć?

Czy da się zmierzyć szczęście linijką?

Pamiętacie, jak kiedyś siedzieliście na lekcji fizyki i uczyliście się przekształcania jednostek miar, i zdecydowanie czuliście, że zostaliście stworzeni do rzeczy większych niż przerabianie cali na metry, że wasze talenty powinny znaleźć zastosowanie w bardziej spektakularnych sytuacjach, na przykład przy hodowli alpaka albo przy czesaniu makaronu? Rozumiem was doskonale, choć w gruncie rzeczy żadna z tych czynności nie jest tak prosta, jak się wydaje, wie to każdy, kto próbował zakręcić loki na makaronie spaghetti (czyli wszyscy). Niemniej w ramach pocieszenia muszę wam powiedzieć, że oto przychodzę do was z ostateczną odpowiedzią na pytanie, dlaczego warto było się tego przekształcania jednostek uczyć. A tej cennej lekcji udzieli nam pewna niewielka firma technologiczna z USA – NASA.

11 grudnia 1998 roku NASA wysłała na Marsa sondę do badań pogody, atmosfery i klimatu. Pracowali nad nią najlepsi inżynierowie z wielu krajów, w tym Wielkiej Brytanii i Stanów Zjednoczonych. Przyznać trzeba, że przez dziewięć miesięcy wszystko szło świetnie, niemniej we wrześniu kolejnego roku, gdy sonda już zbliżała się do celu, utracono z nią kontakt radiowy. Potem było już tylko gorzej – okazało się, że minęła cel, a następnie została zniszczona w atmosferze Marsa. I teraz słuchajcie, jeśli myślicie, że macie zły dzień tylko dlatego, że w drodze do roboty mokry liść przykleił wam się do twarzy, a w sklepie ktoś przypadkiem wrzucił was do beczki z kiszoną kapustą, to pomyślcie sobie o prezesie NASA, który chwilę później musiał wystąpić na konferencji prasowej i wyjaśnić światu, dlaczego misja się nie powiodła. Musiał wystąpić przed dziennikarzami z całego świata i oświadczyć im, że proszę państwa, sytuacja jest taka, że przyczyną był błąd ludzki, być może machnęliśmy się w jednostkach miary, a konkretnie użyliśmy różnych jednostek przy wyznaczaniu trajektorii lotu sondy, albowiem system naziemny został zaprogramowany w anglosaskich funtach, a samo oprogramowanie sondy w metrycznych niutonach i właśnie dlatego te dwa systemy się nie dogadały, a sonda minęła się z celem. Bo w gruncie rzeczy to było zupełnie tak, jakby te dwa systemy mówiły dwoma różnymi językami, a nie że jak człowiek-Polak na wakacjach za granicą, wszędzie się dogada, jeśli tylko będzie mówił po polsku głośno i powoli. Co prawda nie mam pewności, ale wyobrażam

sobie to mniej więcej tak, że gdy ów prezes wygłaszał oświadczenie, to przypominał sobie lekcje fizyki, na których myślał o zaplataniu warkoczy z makaronu spaghetti, i zastanawiał się, do czego mu się to przekształcanie jednostek miar w ogóle w życiu przyda. 20 lat później miał się przekonać, że ta jedna, konkretna lekcja fizyki kosztowała jakieś 327 milionów dolarów.

Wiecie, za 327 milionów dolarów to człowiek mógłby dostać całkiem sporo schabowych w barze mlecznym Renata i jeszcze drugie tyle kompotów, więc jednak zawsze warto zastanowić się, jak odpowiednie własności były mierzone i czy na pewno poprawnie. Przyznać trzeba, że nauki ścisłe mają w tej kwestii trochę prościej, no bo jak taka biologia chce sobie zmierzyć obwód jakiegoś oposa, to po prostu mierzy go w centymetrach, ewentualnie w metrach, jeśli jest to opos, który mieszkał bardzo blisko babci. Jediną trudnością, jaką można napotkać w tym badaniu pomiarowym oposa, jest pewna ruchliwość obiektów badanych, no i przekształcenie jednostek, jeśli zajdzie taka potrzeba⁵¹. Trochę trudniej jest zaś z pomiarem dość abstrakcyjnych pojęć w naukach społecznych, na przykład zadowolenia z pracy, klasy społecznej lub nawet stanu zdrowia. Wie to każdy, kto kiedyś próbował zmierzyć poczucie szczęścia linijką.

W czym zmierzyć szczęście?

Nie, Karol, nie mam tu na myśli pomiarów wykonywanych w wieku nastoletnim w łazience, choć w takich wypadkach korelacja między wynikiem pomiaru a poczuciem zadowolenia rzeczywiście może występować. Gdybyśmy tak naprawdę chcieli się dowiedzieć, jak szczęśliwi są mieszkańcy naszego kraju, to w jaki sposób byśmy to zmierzyli? Wiem, że absolutnie każdy, kto ma w domu puszystokopytnych lokatorów, odpowie teraz, że przecież szczęście można zważyć, wszak on się co rano budzi z czterema kilogramami puchatej radości na twarzy. Niemniej nauki społeczne musiały pomyśleć o trochę innym sposobie pomiaru tego konceptu, głównie dlatego, że każdy, kto ma kota, wie, że owszem, można go zważyć, ale po wykonaniu tej czynności natychmiast trzeba liczyć rannych i zabitych. I właśnie dla ludzi ceniących własne zdrowie i życie oraz dla tych, którzy potrzebują wielu długich słów do pojedynku w scrabble, stworzono pojęcie OPERACJONALIZACJA.

To może ja zacznę od definicji, żebyście wszyscy mogli na jej widok stanąć w płomieniach, a następnie żebym ja mogła to ugasić lodowatym prysznicem edukacji. Operacjonalizacja polega na zdefiniowaniu interesującego nas konceptu, w tym wypadku poczucia szczęścia (jak owo poczucie szczęścia rozumiemy, jaka jest jego definicja, który z mądrych autorów opisał je najlepiej). Następnie zaś na zaprojektowaniu konkretnych operacji, w wyniku których to poczucie szczęścia zmierzymy. Przykładowo moglibyśmy to zrobić tak: przyjmując, że skorzystamy z definicji szczęścia oferowanej przez psychologię pozytywną, która mówi o tym, że „szczęście to stopień, w jakim lubimy swoje własne życie” (Trzebińska 2008). Jeśli przyjmiemy wspomnianą definicję, to poczucie szczęścia możemy zmierzyć na przykład poprzez zadanie respondentom następującego pytania: „Na skali od 1 do 5 jak bardzo jest Pan(i) zadowolony/a ze swojego życia?”. Możemy też zdecydować, że chcemy dokonać bardziej szczegółowego pomiaru dzięki zadaniu innych pytań, na przykład: „Na skali od 1 do 5 jak bardzo jest Pan(i) zadowolony/a ze swojego życia zawodowego?”. A następnie podobne pytania dotyczące pozostałych sfer życia (stanu zdrowia, statusu materialnego lub życia osobistego). I teraz widzicie, my jesteśmy jak kruki, co to za pomocą odpowiedniego narzędzia potrafią dostać się do dżdżownicy w dziurze. W tej metaforze pytanie ankietowe jest narzędziem, a poczucie szczęścia naszą dżdżownicą. Takie narzędzia, których używamy do pomiaru danego konceptu, nazywamy WSKAŹNIKAMI. W tym przypadku wskaźnikiem użytym do pomiaru poczucia szczęścia jest pytanie ankietowe.

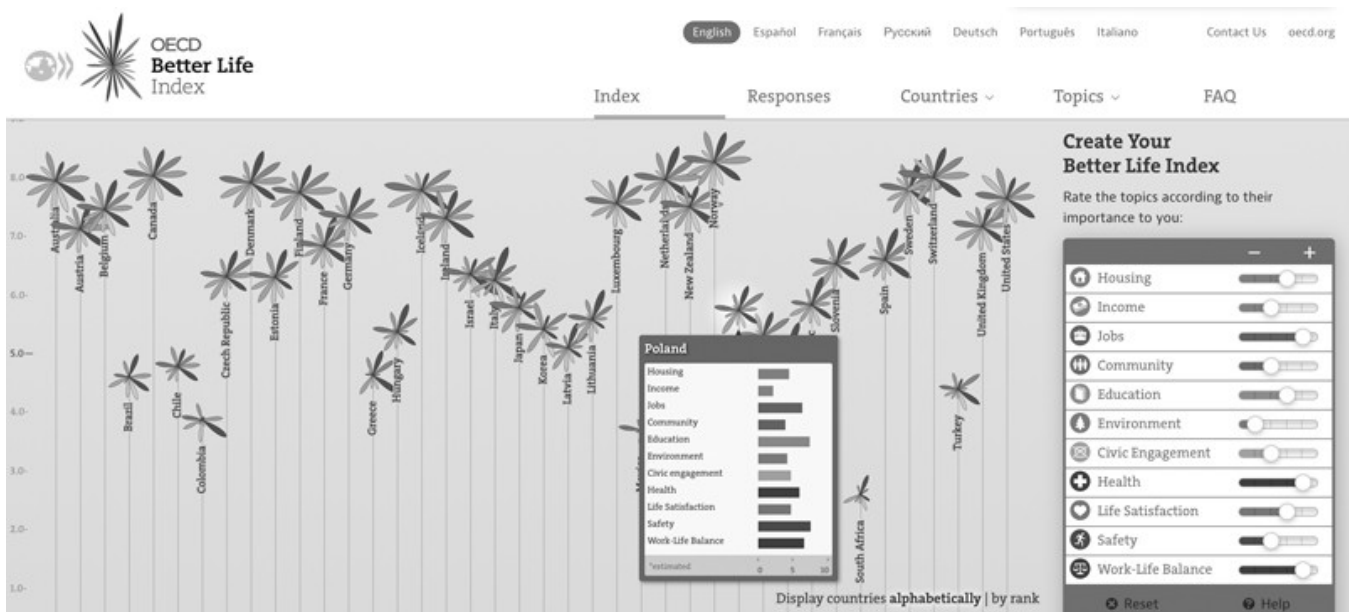
To, jaką definicją się posłużymy, zależy od badacza⁵² i przyjętej przez niego koncepcji badawczej, wybór wskaźników w dużej mierze zależy zaś od przyjętej definicji zjawiska czy problemu. Większość własności można mierzyć na wiele sposobów. Załóżmy, że naszą zmienną jest wytrzymałość fizyczna. Możemy uznać, że wskaźnikiem wytrzymałości fizycznej będzie liczba pompek zrobionych w ciągu minuty. Albo: że będzie to grupa trzech wskaźników składająca się z liczby pompek na minutę, biegu na czas na 100 metrów i pomiaru tętna (lub w moim wypadku – całkowitego braku tętna) po tym ostatnim osiągnięciu. Możemy również potraktować wytrzymałość fizyczną jako subiektywne odczucie – wtedy naszym wskaźnikiem wytrzymałości fizycznej mogłaby być odpowiedź na pytanie ankietowe: „Jak oceniasz swoją wytrzymałość fizyczną na skali od 0 do 5?”. W tym ostatnim wypadku warto jednak zostawić również trochę

miejsca na marginesie, by Janiny⁵³ mogły obok skali dopisać: -10 (ale w 1993 roku zrobiłam przysiad, bo akurat próbowałam zlizać loda, który mi spadł na chodnik).

Skandynawio, co ty bierzesz?!

Poczucie szczęścia to w ogóle fajny przykład, bo rankingów szczęścia mamy na świecie więcej niż szczęśliwych ludzi w naszym kraju ojczystym, no chyba że sąsiadowi akurat kuna zamieszka na strychu albo komuś się truskawki rozsypią w autobusie, bo to zawsze radość. Ale co, jeśli – szalony pomysł – istnieją gdzieś kraje, w których krzywda sąsiada nie cieszy aż tak bardzo?⁵⁴ Albo kraje, w których nie mieszkają kuny, więc nie da się powiązać szczęścia lub nieszczęścia z ich obecnością na strychu? W jaki sposób moglibyśmy porównać poziomy szczęścia w wielu różnych krajach?

Spójrzmy na kilka przykładów. Jednym z bardziej znanych rankingów szczęścia jest OECD *Better Life Index*, czyli narzędzie stworzone po to, by móc porównywać satysfakcję z życia pomiędzy mieszkańcami krajów OECD⁵⁵. W przypadku tego narzędzia satysfakcja z życia mierzona jest za pomocą 11 społecznych wskaźników, takich jak warunki społeczno-bytowe, środowisko, zaangażowanie społeczne, bezpieczeństwo czy spełnienie osobiste.



Screen pochodzi ze strony projektu OECD *Better Life Index*.

Kwestionariusz ankiety zawiera pytania ogólne, w których respondent jest proszony o ocenę na skali od 0 do 10 swojego ogólnego zadowolenia z życia, oraz szczegółowe, dotyczące poszczególnych sfer życia, które według przyjętej definicji składają się na zadowolenie z życia. W 2017 roku średnia wśród badanych krajów wyniosła 6,5; średnia dla Polski – niecałe 5⁵⁶; a na przykład kraje skandynawskie osiągnęły miłe sercu 7,5. No ale wiadomo, tam to każdy byłby zadowolony – wszak oni mają bardzo dużo budulca do lepienia bałwana, a na ich strychach mieszkają niedźwiedzie polarne zamiast kun. Na stronie oecdbetterlifeindex.org znajdziecie interaktywny wykres, na którym możecie sprawdzić, jak przedstawia się szczegółowa punktacja poszczególnych krajów, jak zmieniało się to w czasie, i porównać kraje między sobą, żeby upewnić się, czy jest ktoś, kto ma gorzej od nas. Badanie to przeprowadza się w krajach członkowskich co dwa lata, co pozwala fajnie sprawdzać, w jaki sposób poszczególne kraje stają się szczęśliwsze lub nie.

Jeśli chcielibyśmy porównać deklarowane poczucie szczęścia między krajami, to moglibyśmy spojrzeć również na The World Happiness Report, czyli znacznie większe zestawienie, obejmujące aż 160 krajów. W kwestionariuszu ankiety, która służy do tworzenia raportu (*Gallup World Poll*), można znaleźć pytanie, które również mierzy satysfakcję z życia, niemniej jego autor miał trochę więcej fantazji niż autorzy poprzedniego rankingu. Pytanie znane jest jako drabina Cantrila (*Cantril Ladder*). Zachęca się w nim badanego, by wyobraził sobie drabinę o dziesięciu szczeblach (te szczeble są pięknie ponumerowane, szczyt drabiny reprezentuje najlepsze możliwe życie, a dół – najgorszy z możliwych scenariuszy). Następnie prosi się go o określenie, na którym stopniu drabiny znajduje się obecnie⁵⁷. Według wyników z 2016 roku w tym rankingu ponownie królują kraje skandynawskie, ponownie ze średnią powyżej 7. Co już samo w sobie jest dość irytujące, a to dopiero początek, bo jeśli weźmie się pod uwagę kolejne badanie dotyczące poczucia szczęścia – World Value Survey – to wiecie, ile procent Szwedów twierdzi, że w sumie to są zadowoleni albo bardzo zadowoleni z życia? Ponad 90%! I to się nie zmienia od czasu pierwszej edycji ankiety w 1995 roku!!! Dajcie spokój, tak się kończy to rozpasanie w postaci pracy za godne pieniądze i życia w kraju punktualnych pociągów, a nie że osobowy do Rawicza z godziny 15:12 co prawda przyjedzie o 15:15, ale sześć dni później.

Takich badań, które mierzą poczucie szczęścia w różnych krajach, mamy znacznie więcej i wykorzystują one bardzo różne wskaźniki do pomiaru owej szczęśliwości, czyli tego samego konceptu. Jednym z ciekawszych wskaźników szczęścia, diametralnie różnym od tych dotychczas przedstawionych, jest na przykład światowy indeks szczęścia (ang. *Happy Planet Index*). Jest to miernik ekonomiczny, który ma mierzyć dobrostan w poszczególnych krajach, niemniej w przeciwieństwie do innych wskaźników ekonomicznych, takich jak produkt krajowy brutto czy wskaźnik rozwoju społecznego, bierze on pod uwagę również wpływ na środowisko (tak zwany ślad ekologiczny), czyli czynnik nieobecny w innych badaniach poczucia szczęścia⁵⁸. Żółwie morskie lubią to i na bank klaskałyby z uznaniem, gdyby tylko umiały dosięgnąć płetwą do płetwy.

Okazuje się, że jeśli zmierzmy poczucie szczęścia w ten sposób, to nasi dotychczasowi skandynawscy prymusi zostają daleko w tyle, bo najlepiej radząca sobie Norwegia dotarła ledwie na 29. miejsce. Najszczęśliwymi krajami według tego wskaźnika są zaś Kostaryka, Wietnam i Kolumbia, których próżno szukać w innych rankingach. No i co teraz? Kto nas oszukuje? Kto ma rację?

Ciekawostka: wszyscy ją mają!

Pamiętacie, jak mówiłam wam⁵⁹, że ten sam koncept może być definiowany i operacjonalizowany na wiele różnych sposobów, czyli ostatecznie również różnie mierzony? Siłą rzeczy, gdy definiujemy i mierzymy tę samą rzecz inaczej, to osiągamy nie zawsze zbieżne wyniki. Nie oznacza to, że jeden sposób jest lepszy od drugiego. Oznacza to jednak, że gdy czytacie jakiegokolwiek badanie naukowe, to zawsze powinniście zadawać sobie pytanie: „Jak to zbadano?” z podejrzliwością trzylatka szukającego pietruszki w zupie.

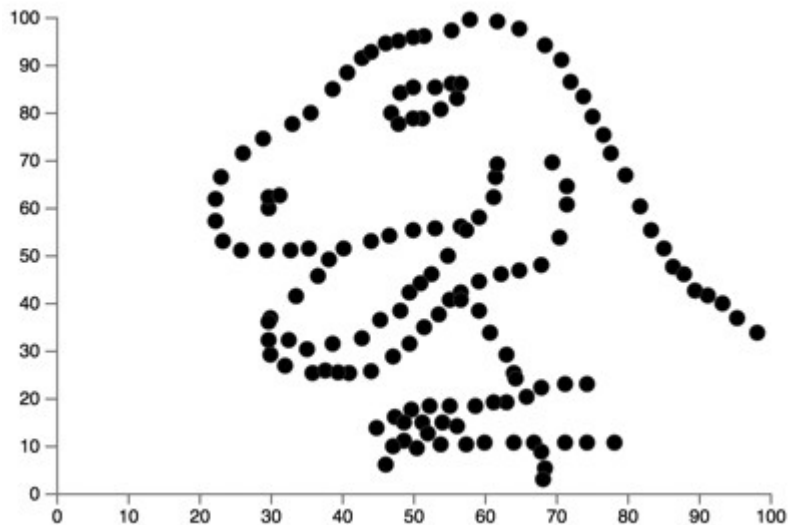
Gdy czytacie jakikolwiek raport lub artykuł naukowy, to zawsze szukajcie paragrafu lub rozdziału, który szczegółowo wyjaśnia, jak badacz zdefiniował, jak zoperacjonalizował i jak zmierzył przedmiot swoich badań. Jest to swoista instrukcja powstawania badania. Pamiętajcie, że sama statystyka bez kontekstu (w tym wypadku wiedzy, jak coś zmierzono) jest jak pierogi bez skwarków lub praca bez przerwy na leżakowanie, czyli bez sensu. A jaranie się wynikami takich badań lub rankingów bez przyjrzenia się ich metodologii to trochę tak jak chwalenie się samochodem ze względu na kolor, bez sprawdzenia, co właściwie ma pod maską. Co w sumie prowadzi mnie do prywatnych wniosków, że potencjał do tego, by

stać się ulubionymi (bo w ogóle niedociekliwymi) klientami wszelkich salonów motoryzacyjnych, mają... przedstawiciele uczelni wyższych.

Gdy twoja uczelnia okazuje się klapkami kubota

Musicie wiedzieć, że mój irlandzki szef kochał rankingi i był bardzo dumny z każdego uniwersyteckiego skoku na rankingowych schodkach, a że pracowałam akurat w Trinity College Dublin, który we wszelkich rankingach radził sobie całkiem nieźle, to mój szef często był szalenie zadowoloną kurą na tym akademickim podwórku. Nawet kazał nam dodawać do maila podpisy brzmiące mniej więcej tak: pozdrawiam, Janina, wydział ten i ten, uniwersytet ten i ten, to miejsce w rankingu światowym, to w europejskim, to w narodowym, i mama też jest dumna. No generalnie w pewnym momencie ten podpis stał się dłuższy niż treść moich maili, zwłaszcza tych wysyłanych do studentów, bo te najczęściej zawierały jeden z dwóch komunikatów, to jest: nie, ewentualnie: chyba żart. To ostatnie zaszło, gdy student wysłał mi wiadomość, że cały poprzedni rok nie dotarł na zajęcia, bo nigdy nie udało mu się wstać na czas, co może nawet spotkałoby się z moim zrozumieniem, gdyby nie to, że zajęcia były o siedemnastej.

No w każdym razie kiedyś poszłam do swojego szefa spytać, czy mogę poprowadzić autorski przedmiot, który nazwałam *Rola zwierząt w statystyce światowej*, i na którym uczyłabym wszystkich, w jaki sposób zaprojektować badania naukowe z owcami (Kendrick et al. 2001) oraz w jaki sposób narysować dinozaura z danych (Matejka i Fitzmaurice 2017), o takiego:



Wykres powstał na podstawie danych udostępnionych w ramach projektu The Datasaurus Dozen.

I jeszcze całość nauki wieńczyłaby taka uroczysta gala, podczas której prezentowalibyśmy wszystkim swój hymn o roli zwierzątek w statystyce, z kulminacyjnym momentem w postaci refrenu o tym, że:
„Kot ma uszy, koza róg /Pies ze mną na spacerze nie ma trzech nóg”.

I jeszcze w tle takie chórkiki: „uuuu”, bo wiadomo, że jak jest chórek, to jest profesjonalizm.

No więc poszłam przedstawić szefowi swoją koncepcję, a on wysłuchał jej z uwagą, a następnie rzekł: Janina, co ty bierzesz?! Choć może to dlatego, że nie miałam chórków, co ujęło nieco efektywności refrenowi. No i dlatego, że mój szef był wtedy najsmutniejszym człowiekiem na świecie, bo od rana gazety krzyczały, że w tym roku żadna irlandzka uczelnia nie znalazła się w setce najlepszych na świecie, mimo iż przez długi czas trzymaliśmy się pazurami pierwszego miejsca we wszelkich krajowych rankingach i minimum pierwszej setki w światowych zestawieniach. Jak to się stało, spytacie. Czy w ciągu roku zdołaliśmy aż tak popikować w dół, niczym rybitwa, która dojrzała śledzia? Otóż odpowiadam: w roku, o którym mowa, w ogóle nie było nas w rankingu, bo okazało się, że w informacjach, które wysłaliśmy jego twórcom, ktoś się pomylił podczas obliczeń i źle wstawił przecinek w danych dotyczących przychodu netto. Czyli generalnie okazało się, że nie potrafimy liczyć. A jak już człowiek myślał, że gorzej być nie może, to okazało się, że przez dwa lata z rzędu liczyliśmy wszystko źle, więc nas zdyskwalifikowano wstecz, i ja uważam,

że to przezabawna historia, i przez trzy dni się śmiałam, gdy ją usłyszałam, ale mój szef się w ogóle nie śmiał, on tylko strasznie wzdychał, a im bardziej on wzdychał, tym bardziej ja się śmiałam, i on jeszcze czasem zawieszał na mnie wzrok, wzdychał i słabym głosem pytał mnie: *Janina, why?*, a ja sobie wtedy myślałam, że jeśli ten uniwersytet przez dwa lata nie nauczył się, jak prawidłowo obliczać swój przychód netto, to może to wiele mówi o naszej pozycji w rankingu, ale nic nie mówiłam na głos, bo w obliczu rozpaczy mojego szefa to by było jak kopanie kaczątka i oblewanie go wodą ze szlauchu.

Nie no, serio, czytaliście kiedyś jakiś ranking uczelni wyższych? To dopiero jest zabawa, gdy nagle okazuje się, że ta sama uczelnia w jednym tygodniu chwali się pierwszym miejscem po to, by za chwilę w innym rankingu być już daleko poza podium. Wynika to oczywiście z różnic w przyjętej metodologii i jest tak samo naturalne jak podstępne. Obecnie jest tak wiele rankingów, że jeśli chcecie podjąć decyzję o tym, gdzie studiować, na podstawie takich zestawień, to warto sprawdzić, za co właściwie przyznano uczelni punkty i czy przypadkiem nie za różnorodność batoników czekoladowych dostępnych w automatach na korytarzu. Bo o ile ten wskaźnik może pozytywnie wpływać na waszą jakość życia, to na jakość kształcenia niekoniecznie.

Z wszelakimi rankingami jest podobny problem jak z mikołajami z czekolady znalezionymi w świątecznych paczkach – dopóki nie odgryziesz mu głowy, dopóty nie dowiesz się, czy jest to prawdziwa czekolada, czy wyrób czekoladopodobny. Pozostańmy przy tej metaforze – chociaż gazetowe nagłówki często kuszą nas doniesieniami, że „Wyższa Szkoła Wszystkiego na trzecim miejscu najlepszych uczelni wyższych w powiecie zielonogórskim” (z trzech) czy „Polskie psy na ostatnim miejscu w rankingu najbardziej puchatych psów w Europie”, to zawsze powinniśmy podchodzić do nich z pewnym sceptycyzmem. Zanim wmasujemy w naszego burka oliwkę zwiększającą objętość sierści, pomyślmy przez chwilę, na ile dany ranking jest wiarygodny.

Dla przykładu w Hiszpanii stworzono ranking uniwersytetów (<http://www.webometrics.info>), który dość swobodnie podchodził do kwestii jakości kształcenia czy oferty edukacyjnej, bo najbardziej zainteresowany był liczbą i objętością tekstów opublikowanych na stronach internetowych uczelni, ich przejrzystością i częstotliwością odwiedzin. No i w porządku, bo faktycznie może być tak, że te teksty na stronie

internetowej są dla kogoś kluczowe, że ktoś już sobie totalnie wyobraża, jak pójdzie na rozmowę o pracę i go spytają, który element kształcenia uniwersyteckiego był dla niego najważniejszy w procesie wzrostu osobistego, a on wtedy wreszcie będzie mógł powiedzieć, że generalnie to najbardziej wzbogacił go wewnątrz ten tekst w dziale Aktualności dotyczący remontu stołówki, bo po zmianie koloru ściany będą zielone jak brukselka na tle zachodzącego słońca.

Jeśli chodzi o inne zestawienia uczelni wyższych, to moim ulubionym wciąż pozostaje ranking zwany The Campus Squirrel Listings, który porządkuje amerykańskie uniwersytety, stosując kryteria takie jak wielkość, zdrowie i zadowolenie wiewiórek żyjących na danym kampusie. No i umówmy się – to są ważne rzeczy. Jak wiewiórka Elżbieta zastanawia się, do którego akademickiego koszyczka włożyć swoje orzeszki, to ten ranking będzie dla niej szalenie istotnym źródłem informacji.



Zrozumiałe jest pragnienie wszelkich autorów i instytucji, by stworzyć narodowy lub międzynarodowy ranking uczelni, który jasno wskaże zainteresowanym, że tu jest fajnie, a tam jest niefajnie. Ale jest to zadanie szalenie trudne, biorąc pod uwagę, jak różne są narodowe systemy edukacji, ich struktura, finansowanie i możliwości. To znaczy – trudno porównywać Harvard, jedną z najlepszych i najbogatszych uczelni na świecie, o rocznym budżecie od 32 do 38 miliardów dolarów (zależnie od źródła), z Kabul Medical University w Afganistanie, który najpewniej przegrywa z Harvardem na wielu frontach (na przykład liczby stypendystów czy oferty kursów), niemniej jest jednym z niewielu uniwersytetów w swoim kraju, które umożliwiają kobietom bezpieczną edukację wyższą dzięki wprowadzeniu udogodnień, takich jak żeńskie łazienki (tak jest, nie we wszystkich krajach świata jest to oczywiste!) lub

autobusy bezpiecznie dowożące studentki na uczelnię. I nawet jeśli ten uniwersytet nie może się pochwalić laureatką czy laureatem Nagrody Nobla, to nie da się przecenić jego roli w zapewnianiu równego dostępu do edukacji. Mówimy przecież o kraju, w którym jeszcze w 2002 roku kobiety stanowiły 13% studenckiej populacji, a w 2012 roku ta liczba owszem, wzrosła, ale do oszałamiających 20% (Samady 2013). Jasne więc, że ta uczelnia nie zdobędzie zbyt wielu punktów w kategorii międzynarodowych publikacji czy wymian studenckich, niemniej w tamtej rzeczywistości społeczno-kulturowej jest szalenie istotną instytucją.

Innymi słowy, chciałam was zachęcić do tego, by w zestawieniach – czy to uczelni wyższych, czy najlepszych pracodawców, czy też najgroźniejszych ras psów – zawsze sprawdzać, w jaki sposób dane wskaźniki zmierzono. I zawsze myśleć o tych rankingach trochę szerzej, krytycznie. W gruncie rzeczy to, czy pędzimy autostradą sukcesu fiatem 126p, czy też innym środkiem lokomocji, ma spore znaczenie.

Fiat 126p edukacji uniwersyteckiej czy Pendolino szkolnictwa wyższego?

No to spójrzmy na przykład trochę bardziej swojski, rzekłabym – takie pierogi pedagogicznego świata, a nie egzotyczne *gholebi palaw* afgańskiej edukacji. Spójrzmy na dwa polskie rankingi szkół wyższych najczęściej cytowane przez media – ranking szkół wyższych „Polityki” i „Perspektyw”.

Gdy przyjrzymy się kryteriom, które są brane pod uwagę w rankingu „Perspektyw”, to znajdziemy tam takie wskaźniki jak PRESTIŻ (mierzony oceną kadry akademickiej i pozycją w rankingach międzynarodowych), będący jednym z moich ulubionych kryteriów i mówię to z całą świadomością człowieka, którego ukończył UJ (bo ujotu się nie kończy, to ugot kończy ciebie), na którym edukacja jedynie w połowie składała się z zajęć dydaktycznych, bo przez pozostały czas byliśmy zajęci płakaniem ze wzruszenia przed portretami sławnych absolwentów naszej uczelni, czytaniem książki *Rola Uniwersytetu Jagiellońskiego w dziejach narodu Polskiego* (serio!!!) i słuchaniem o tym, że jesteśmy najpiękniej wyrosniętymi stokrotkami na tej narodowej polanie pełnej szczawiu. Nie winię jednak „Perspektyw” za to, że biorą pod uwagę to kryterium w swoim rankingu, bo tak robi wiele światowych rankingów i spoko, można na to kryterium spojrzeć z pewną ciekawością, niemniej warto mieć w pamięci, że jak już człowiek tę swoją Alma Mater skończy, to kanapką

z prestiżem raczej się nie naje, a na dodatek pani w gazowni uparcie twierdzi, że estymą nie można płacić rachunków.

Innym kryterium brany pod uwagę w rankingu „Perspektyw” są absolwenci na rynku pracy, które to kryterium jest mierzone za pomocą dwóch wskaźników. Po pierwsze, za pomocą wskaźnika PREFERENCJE PRACODAWCÓW, który określa, ilu z ankietowanych pracodawców powiedziało, że CV absolwentów danej uczelni świeci dla nich jakby jaśniejszym blaskiem, pachnie świeżo pieczonym ciastem drożdżowym i jeszcze przy przekładaniu stron słychać w tle ciche *Ave Maria* przygrywane na skrzypcach. Po drugie, uwzględniany jest taki wskaźnik jak EKONOMICZNE LOSY ABSOLWENTÓW, czyli ilu absolwentów znalazło zatrudnienie po studiach i jaka jest wysokość ich zarobków. I ten wskaźnik jest moim zdaniem szalenie istotny, i fajnie, że się go mierzy, zwłaszcza jeśli ktoś nie skończył Uniwersytetu Jagiellońskiego i nie może sobie zająć na lunch sałatki ze splendoru polanej sosem zrobionym z autorytetu i poważania innych ludzi.

A mieliście na studiach tego jednego człowieka, który w samym środku zawodów w ubijaniu bitej śmietany głową, odbywających się na ostatnim piętrze akademika Fistaszek, mówił: „Ej, dziewczyny i chłopaki, chyba już czas przerwać tę niepoważną zabawę i iść spać, wszak jest już 21:15, a jutro kolokwium”? No właśnie. To takich ludzi na bank zainteresuje, czy ów ranking w jakikolwiek sposób mierzy to, czy można się na tych studiach czegokolwiek nauczyć, to znaczy – oprócz tego, jak przetrwać miesiąc za dwa złote osiemdziesiąt pięć groszy.

I owszem. W rankingu „Perspektyw” znajdziemy też takie parametry jak EFEKTYWNOŚĆ NAUKOWA i POTENCJAŁ NAUKOWY, które są mierzone między innymi takimi wskaźnikami: nadane stopnie naukowe, uprawnienia doktorskie, uprawnienia habilitacyjne, publikacje, cytowania i tym podobne. Jakbyście chcieli sprawę uprościć, to są to takie wskaźniki, które mają mierzyć jakość kadry, czyli czy uczą nas przyszli nobliści, czy też taka Janina, która doktorat robi tyle lat, że potem to już się nie umiała doliczyć, ile ich minęło, co nie byłoby nawet takie złe, gdyby nie to, że doktorat robi akurat z liczenia. To ważne wskaźniki, bo pokazują potencjał kadry naukowej, właściwie można powiedzieć, że szacują prawdopodobieństwo ich wygranej w 1 z 10, ewentualnie w *Milionerach*⁶⁰. Niemniej mam tutaj jedno ważne zastrzeżenie.

O ile liczba publikacji międzynarodowych, cytowań w wysoko punktowanych czasopismach i tytułów naukowych może w jakiś sposób świadczyć o poziomie badań prowadzonych w danej jednostce, o tyle niekoniecznie przekłada się w jakikolwiek sposób na jakość kształcenia i same doświadczenia studentów. W gruncie rzeczy wszyscy mieliśmy na studiach takiego profesora, który był wybitnym ekspertem we własnej dziedzinie, ale jeśli chodzi o umiejętność przekazywania wiedzy, to szło mu to tak, rzekłabym, nienachalnie. Bo na wykładach najczęściej czytał z kartek, a jak mu się dwie strony skleiły, to często nawet tego nie zauważał, tworzył przez to zdania zupełnie bez sensu, a poza tym slajdy na rzutniku puszczał do góry nogami i wszystkie wykłady kończył wcześniej, by móc w tym czasie biegać bosą po łące, tańczyć w deszczu i pleść wianki ze stokrotek.

To sprawia, że o rankingu naukowym uczelni akademickich publikowanym corocznie w „Polityce” musimy myśleć ostrożnie, albowiem tam klasyfikacja jest tworzona na podstawie INDEKSU HIRSCHA, na który składa się liczba publikacji naukowych powstałych w danym ośrodku i istotność owych publikacji mierzona liczbą cytowań. Oprócz wątpliwości podobnych do tych, o których wspominałam w kontekście wybitnego naukowca, który nie miał czasu uczyć studentów, bo był zajęty łapaniem motyli i obserwacją traszek, sami autorzy rankingu zwracają uwagę na dodatkowe problemy związane z takim systemem tworzenia klasyfikacji, wśród których kluczowe jest to, że indeksu Hirscha nie da się zastosować w naukach humanistycznych, ponieważ w bazach używanych do jego przeliczenia niewiele jest czasopism humanistycznych. Wynika to z tego, że są to bazy stworzone z myślą o naukach ścisłych, medycznych i przyrodniczych, do pewnego stopnia również naukach społecznych (choć to zależy od dyscypliny). Widzicie więc, że to ograniczenie wydaje się dość znaczące, albowiem taki ranking nie bierze pod uwagę mnóstwa wydziałów, kierunków, dyscyplin i naukowców. To trochę tak, jakbyście poszli na spacer do lasu z grupą dwadzieścioro przedszkolaków, a po powrocie doliczyli się jedynie połowy i uznali, że spoko, jest to akceptowalny błąd pomiaru. Hubert z grupy Muchomorków, który od trzech lat błąka się po lesie i poluje na ryjówki, najpewniej ma inne zdanie na ten temat.

Czy to znaczy, że ranking „Perspektyw” jest lepszy od tego tworzony przez „Politykę”? A może istnieje jeszcze jakieś inne zestawienie, któremu

można ufać bardziej? Nie. To oznacza po prostu, że przy czytaniu jakiegokolwiek rankingu musimy zwracać uwagę na metodologię i najpierw doczytać, co i jak zostało zmierzone. Musimy również mieć w pamięci, że niektóre wskaźniki są mniej wiarygodne niż inne.

Wróćmy na przykład do mojego ulubionego wskaźnika, czyli prestiżu lub reputacji – o ile na pierwszy rzut oka może nam się wydawać intuicyjny, to w gruncie rzeczy jest bardzo problematyczny. Po pierwsze dlatego, że jest szalenie niejednoznaczny – dla każdego z respondentów reputacja może oznaczać coś innego, czyli być inaczej rozumiana czy definiowana. Po drugie dlatego, że ten wskaźnik jest w dużej mierze obarczony błędem atrybucji zwanym EFEKTEM HALO, który oznacza, że mamy automatyczną tendencję do przypisywania komuś lub czemuś pozytywnych lub negatywnych właściwości na podstawie pierwszego wrażenia, nawet gdy tych właściwości nie zaobserwowaliśmy. To znaczy, że nasza ocena reputacji będzie wynikać raczej z dotychczasowych przekonań na temat prestiżu danej placówki niż z systematycznej i krytycznej ewaluacji jakości programów dydaktycznych. Chciałabym myśleć, że za każdym razem, gdy ktoś nas prosi o ocenę reputacji danego uniwersytetu, podejmujemy decyzję na podstawie analizy fenomenologicznej programów kształcenia i zaawansowanej hermeneutycznej interpretacji tekstów źródłowych. Niemniej może być trochę inaczej i okaże się, że takie decyzje podejmujemy, bazując na tym, co już kiedyś słyszeliśmy, co kiedyś ktoś nam powiedział, ewentualnie – ile razy napatrzyliśmy się na portrety Mikołaja Kopernika i Jana Pawła II w czasie inauguracji roku akademickiego.

Inną istotną kwestią dotyczącą rankingów, na którą musimy zwracać uwagę, jest to, czy dane dotyczące klasyfikowanych uczelni są zbierane przez niezależne podmioty, czy też dostarczane przez samych zainteresowanych. Na przykład o ile w Wielkiej Brytanii i Australii to agencje rządowe zbierają i analizują dane dotyczące uczelni, o tyle w Stanach Zjednoczonych bardzo często to same uczelnie są zobowiązane do przedstawienia takich danych. Czyli to trochę tak, jakby dać uczniom dziennik i kazać im samym wpisać swoje oceny z ostatnich kartków. W takim wypadku – parafrazując słowa ze słynnej gry komputerowej – najpewniej doszłoby do sytuacji, w której populacja geniuszy zwiększa się.

Nie twierdzę więc, że wszystkie rankingi – wliczając w to te dotyczące szkół podstawowych, liceów lub uczelni wyższych – są do bani.

Powinniśmy jednak przyglądać się im z uwagą i dokładnie sprawdzać, jak niektóre wskaźniki zostały zmierzone, by móc na ich podstawie wyciągać jak najlepsze wnioski. I zawsze zadawać sobie pytanie, czy ów ranking i przyjęte przez niego wskaźniki naprawdę mierzą efektywność danej instytucji, czy raczej jej efektywność. Zwłaszcza że rankingi to nie największe wyzwanie, z którym przychodzi się zmierzyć metodologom (he, he, czaicie? zmierzyć!). Bo w rankingach to jeszcze mamy do czynienia z danymi zastanymi, którym trudno odmówić obiektywizmu: liczba publikacji, liczba profesorów, liczba studentów zagranicznych. Ale co się dzieje wtedy, gdy to, co chcemy zmierzyć, wymaga pracy z kruszczem delikatnym jak sernik? Wiecie, INNYMI LUDŹMI? Zastanówmy się, czy to w ogóle ma sens. Czy ankietom można ufać? Lub też – jeśli to poprzednie pytanie nie sprawiło, że podjaraliście się jak lampion w roraty – ujmę sprawę inaczej: czy po seksie będziemy opanierowani w banknotach niczym krokiet w bułce?

Czy po seksie będziemy opanierowani w banknotach niczym krokiet w bułce?

Czasem chciałabym być nauczycielem geografii, bo wtedy mogłabym usiąść na środku sali i kazać studentom biegać dookoła mnie, i tłumaczyć im, że ja jestem słońcem, oni planetami, a tak właśnie wygląda Układ Słoneczny. Oczywiście nie narzekam, bo przedmioty, których uczę, też dają mi spore możliwości, no ale ileż można przychodzić na zajęcia przebranym za kalkulator.

Chociaż pochwalę się wam, że kiedyś wczołgałam się trochę na pagórek kreatywności, gdy tłumaczyłam moim studentom, że średnia arytmetyczna to taki sprzedajny labrador miar tendencji centralnych, czyli niby miły, niegroźny, wszyscy go lubią, ale tak naprawdę szalenie łatwo daje się zmanipulować kiełbasą wartości skrajnych. I opowiadałam na przykład, że gdybyśmy chcieli policzyć średnią masę ciała w naszej grupie, no to jest to dość proste, dodajemy: ja – 50 kilogramów (tu śmieję się jak opętana z własnego dowcipu), później waga kolegi, koleżanki, dzielimy przez liczbę osób, ciach, jest średnia, i to całkiem rzetelna, no bo wszyscy ważymy podobnie. Ale co by było – tłumaczę dalej – gdybyśmy chcieli policzyć średnią masę ciała wszystkich w tej sali, a tu by nagle weszło troje zawodników sumo, i to nie że tacy zwykli zawodnicy, ale tacy wiele lat po ślubie, czyli jeszcze bardziej okrągli, i oni przychodzą i siadają w ławce, i co my byśmy wtedy zrobili?

A oni mówią: zdjęcia.

No ja akurat miałam na myśli liczenie mediany, ale za wrzucenie mediany na Fejsa to się jednak dostaje znacznie mniej lajków, wiem, bo sprawdzałam. Widzicie, to jest właśnie moje życie. A w tym samym momencie gdzieś tam jakiś nauczyciel geografii świetnie się bawi, udając Układ Słoneczny.

Niemniej jeśli chodzi o dobre wiadomości, to mam taki jeden temat z metodologii, który bardzo lubię i w którego uczeniu przeszłam sama siebie. Po tych zajęciach biję sobie brawo i wysyłam do siebie kartki gratulacyjne, a że irlandzka poczta się nie śpieszy, to zapominam, że je wysłałam, i jak one później do mnie przychodzą, to się strasznie jaram, że ktoś wysłał do mnie kartkę i jeszcze na dodatek mówi mi w niej, że jestem

świetna. To są takie zajęcia, na potrzeby których stworzyłam turniej o badaniach ankietowych, mój Boże, to są takie zabawne rozgrywki, że to tylko kwestia czasu, aż kupi go ode mnie jakaś telewizja, a prowadzić go będzie Małgorzata Rozenek, bo trochę mam nadzieję, że jak przyjdzie podpisywać kontrakt, to mi przy okazji posprząta.

Muszę wam powiedzieć, że kocham badania ankietowe, to mój ulubiony temat. Swoją córkę nazwę Kafeteria, uszyję jej pluszaka w kształcie kwestionariusza, a potem będziemy się wspólnie bawić w operacjonalizację zmiennych, przynajmniej do czasu, aż dorośnie na tyle, że zadzwoni do odpowiednich instytucji, by powiadomić je, że sama chce się oddać do adopcji.

W każdym razie moja zabawa polega na tym, że mamy drużyny i różne pytania, i zadania, i próby przekupstwa mojej osoby, i wszyscy bardzo dużo krzyczą, i generalnie moi studenci są wtedy jak takie entuzjastyczne koty, które nie do końca wiedzą, że w miłym pasztecie z dzika tak naprawdę podtykam im gorzką tabletkę edukacji. No więc gramy, są emocje, kolejna runda, pytanie dotyczy EFEKTU ANKIETERA (ang. *interviewer effect*), czyli tendencji do fałszowania odpowiedzi ze względu na sposób zachowania lub wygląd ankietera. W tym miejscu zawsze podaję przykład, że to trochę tak, jakby wysłać panią kurę po domach, żeby przeprowadziła wśród ludzi ankietę, czy lubią rosół, i spodziewać się, że ludzie będą odpowiadać zgodnie z prawdą.

No ale wróćmy do turnieju. Sytuacja wygląda tak, że ja pytam, jak nazywamy taką tendencję do odpowiadania na pytania tak, żeby przypodobać się ankieterowi, a oni w ogóle nie potrafią sobie z tym pytaniem poradzić. Myślą, myślą, długo nie wiedzą, więc im podpowiadam, że *inter... intervieweeeeeeer...* i wtedy jeden przypomniał sobie, ryknął: *INTERVIEWER EFFECT!!!!!!!!!!*, a to była poprawna odpowiedź, POPRAWNA ODPOWIEDŹ, więc w mojej głowie to brzmiało tak, jakby właśnie anioł przygrywał *Odę do radości* na harfie, co ma jedwabne struny plecione przez łabędzie, liście spadające z drzew zatrzymały się z tego zachwyty w pół drogi, wszystkie delfiny świata wyskoczyły z wody, ułożywszy się w kształt ogromnego serca, i gdzieś tam na świecie narodziły się małe labradory.

A po skończonych zajęciach powiedziałam, że zwycięska grupa może wybrać sobie nagrodę – albo na następnych zajęciach będziemy jeść cukierki, albo skończymy wykład 10 minut wcześniej, a oni powiedzieli, że

wolą cukierki, co mnie w sumie trochę zmartwiło, bo ja mam dopiero trzydzieści lat i jeszcze nie umiem się dzielić z innymi dziećmi. Potem powiedziałam jeszcze koledze od ostatniego pytania, że może sobie wybrać nagrodę specjalną, on zapytał jaką, a ja na to, że w sumie dowolną, a on pyta, czy mam ksero w biurze, powiedziałam, że mam, a on, że to wspaniale się składa, bo z wszystkich rzeczy na świecie to on już od wielu lat ma takie marzenie, tylko jedno marzenie, i zapamiętajcie ten moment, bo być może i wam kiedyś przyjdzie skorzystać z tej czułości losu, jaką jest spełnienie najskrytszych pragnień zapisanych po wewnętrznej stronie waszych serc. Być może też kiedyś staniecie przed tak ważnym wyborem i będziecie wiedzieć, że oto jest ten najważniejszy moment w waszym życiu, którego nie wolno zmarnować, albowiem – jak mawiał Walt Disney – „czymże bylibyśmy bez marzeń? Ziemią jałową, co na deszcz ożywczy z utęsknieniem czeka”.

Jedno takie wielkie marzenie mam, Janina – zdradził mi student – chciałbym sobie skserować twarz.

No, w gruncie rzeczy dobrze, że twarz.

I to nie jest ostatnia dobra wiadomość, bo jeśli i wy macie życiowe marzenia o takiej randze, to oto przedstawiam wam rozdział, po którego przeczytaniu będziecie w stanie wygrać każdy turniej badań ankietowych, choć w sumie nie będzie to w ogóle konieczne, bo przy okazji zdradzę wam sposób na to, by opanierować się w pieniądzu jak krokiet w bułce.

Jak zasłużyć na ksero własnej twarzy?

2007 rok to był w ogóle ciekawy rok. Czy wiecie na przykład, że był to rok ogłoszony rokiem Bolesława Leśmiana, Stanisława Wyspiańskiego, Artura Rubinsteina, ale też... delfina? W sensie takiego zwierzęcia. Rany, jego mama musiała być dumna, swój własny rok dostał i przez dwanaście miesięcy wszystkie rozgwyzdy chodziły w T-shirtach z jego podobizną, a płaszczki sprzedawały papierowe wiatraczki na festynach organizowanych na delfinią cześć. W 2007 roku wydarzyło się coś jeszcze – w Niemczech przeprowadzono badanie, którego wyniki okazały się szalenie obiecujące dla bardzo wielu z nas. Było to bardzo duże badanie ankietowe na próbie 7500 domostw. Pytano domowników o najróżniejsze rzeczy – jak dojeżdżają do pracy, jak spędzają czas wolny, jakie są ich ulubione programy telewizyjne... no możecie sobie wyobrazić, że te pytania były

trochę jak bigos cioci Heli na ostatnim weselu kuzyna – absolutny miszmasz i bałagan totalny, było tam absolutnie wszystko i nie możemy wykluczyć, że zaginiony chomik sześćioletniej Marlenki też. Przewaga tego badania nad bigosem cioci Heli była jednak oczywista. Po pierwsze, najpewniej w procesie jego tworzenia nie zginął (prawdopodobnie) żaden gryzoń. Po drugie, dzięki zebraniu tak dużej liczby danych i tak różnorodnych zmiennych można było pobawić się w szukanie związków i zależności pomiędzy czynnikami, które niekoniecznie wydawały się ze sobą powiązane. Tak też zresztą się stało – badacze skorelowali wiele różnych czynników i ze zdumieniem odkryli zależność, która sprawiła, że badacz Hans natychmiast poszedł odkopać swój karnet na siłownię, a badacz Ulrich aż odłożył nadgryzione ciastko. Okazało się mianowicie, że ludzie, którzy uprawiają seks 3–4 razy w tygodniu...

(tutaj potrzymam was chwilę w niepewności)

...zarabiają o 3,4% więcej niż pozostali.

No dajcie spokój, każdy by się podjarał, zawsze to prostszy sposób na zostanie miliarderem niż systematyczna praca zawodowa połączona z ograniczonym wydawaniem pieniędzy. Nic więc dziwnego, że to odkrycie natychmiast zostało podchwyczone przez gazety, które ogłosiły: ej, uprawiajcie więcej seksu, by lepiej zarabiać!

Musicie wiedzieć, że gdy podaję ten przykład na swoich wykładach, to w tym miejscu robię strategiczną przerwę, bo połowa moich studentów jest już za drzwiami. Już są w drodze do akademika w trosce o swoją finansową przyszłość. Wystarczyłoby jednak, żeby chwilę poczekali, a zrujnowałabym im życie. Z rozdziału *Ile trzeba zjeść czekolady, żeby dostać Nobla?* dowiemy się, że sam związek pomiędzy dwoma czynnikami nie oznacza automatycznie, że jedno powoduje drugie. Zastanówmy się więc, jaki może być związek między seksem a zarobkami?

Może być tak, że im więcej seksu, tym większe zarobki. Że jak Julian ma dobrze w chałupie, to mu się do tego domu śpieszy, pracuje sto razy bardziej efektywnie i przez to jest obsypywany złotem i awansami. Może tak być.

Opcja druga jest odwrotna: im większe zarobki, tym więcej seksu. Może być tak, że jak Julian wraca do domu i od progu krzyczy, że jest milionerem, to Halinie płoną lędźwia. No tak też może być.

Opcja kolejna jest znacznie mniej spektakularna i polega na tym, że istnieje jakaś trzecia zmienna, która wyjaśnia związek między seksem

i zarobkami. Na przykład stan zdrowia lub określone cechy osobowości, które powodują, że ludzie, którzy je mają, więcej zarabiają i dlatego ich życie seksualne jest urozmaicone jak u mątwy w czasie godów. Nie no, serio, wiedzieliście, że mątwy w czasie godów zmieniają kolory?!

I opcja ostatnia: istnieje inne wyjaśnienie tej korelacji. Wliczając w to przypadek, bo takie przypadkowe korelacje też się zdarzają. Na przykład nie wiem, czy wiecie, ale istnieje bardzo silna korelacja między liczbą filmów, w których grał Nicolas Cage, a liczbą utonień w basenie. I choć każdy, kto widział go w jakimkolwiek filmie, jest w stanie uwierzyć w nagłe myśli samobójcze po projekcji, to rozwiązanie tej zagadki jest znacznie prostsze i brzmi: losowość. Choć oczywiście przezorny zawsze ubezpieczony, więc dobrze by było, gdybyście nie oglądali żadnego filmu z Nicolasem Cage'em. W sensie, wiecie, nie ma co ryzykować.

No to wróćmy do tego seksu i zarobków. Zanim podjaracie się tym odkryciem jak jeź na wykładach akupunktury, wytłumaczę wam, że wyjaśnieniem tej zależności jest punkt ostatni, czyli: istnieje inne uzasadnienie wspomnianej korelacji. Mianowicie – zapomnieliście, że ludzie to kłamczuszki. I że istnieje w badaniach ankietowych coś takiego, co nazywamy EFEKTEM SPOŁECZNYCH OCZEKIWAŃ. To znaczy, że na niektóre pytania odpowiadamy niekoniecznie zgodnie ze stanem faktycznym, ale tak jak (wedle naszego mniemania) wygląda oczekiwanie społeczne. To generalnie jest dość ciekawe, bo badania pokazują, że są zachowania, których częstotliwość lubimy zawyżać (na przykład chodzenie na siłownię, ale o tym to tylko słyszałam, bo ja akurat na to pytanie zawsze odpowiadam zgodnie z prawdą, że na siłowni byłam trzy razy, z tym że być może nie dopowiadam, że trzy razy w życiu, a nie że trzy razy w tym tygodniu), i są też takie czynności, których prawdziwą częstotliwość zaniżamy, na przykład korzystanie z wszelkich używek typu alkohol, papierosy, a w przypadku mojego męża – również spożycie kiełbasy. Wyjątkiem jest częstotliwość (niepełnoletni niech zasłonią teraz oczy) kontaktów seksualnych, a tym, którzy nie wiedzą, tłumaczę, że kontakt seksualny jest wtedy, kiedy osoba z osobą się bardzo kochają i pewnego dnia dają sobie buzi na golasa, i to jest miłe, chyba że są przed ślubem, to niezbyt miłe, bo spłoną w piekle. No i z tą liczbą partnerów i kontaktów seksualnych jest tak, że mężczyźni mają tendencję do jej zawyżania, a kobiety do zaniżania. Ciekawe, nie? Mam nadzieję, że w tym momencie choć pięcioro z was wstało i zaczęło klaskać z uznaniem.

Nie jest więc prawdą, że uprawiając więcej seksu, zarobimy więcej. Prawdą jest natomiast, że mężczyźni z wielu różnych powodów wciąż zarabiają więcej od kobiet. I uprawiają więcej seksu. A przynajmniej tak deklarują. Klawa odpowiedź na zagadkę, nie? A to nie jedyna sytuacja, kiedy lubimy w badaniach ankietowych dostosowywać nasze odpowiedzi do cudzych upodobań.

Czy powiesz to tej krowie prosto w pysk?

EFEKT SPOŁECZNYCH OCZEKIWAŃ to niestety nie jedyne utrudnienie, które może nas czekać w przypadku przeprowadzania badań ankietowych. Możemy mieć również do czynienia z wyżej już wspomnianym EFEKTEM ANKIETERA, czyli sytuacją, w której charakterystyka osoby zadającej pytania – na przykład jej rasa lub płeć – może wpływać na sposób odpowiedzi. Znaczący wyobraźcie sobie, że przychodzi do was jakaś krowa, puka do drzwi kopytkiem, puk, puk, i mówi: „Dzień dobry, przeprowadzam tu takie badanie, czy mógłby mi Pan powiedzieć, jak często jada Pan hamburgery?” No i weźcie powiedźcie teraz tej krowie prosto w pysk, że w sumie to po kilka razy w tygodniu macie rendez-vous z kotлетem i być może jesteście również skarbnikiem w Stowarzyszeniu Miłośników Mielonego Mięsa⁶¹. Zresztą to nie tylko o krowy chodzi, ale o wszelkie badania, w których cechy demograficzne ankietera mogą w jakikolwiek sposób wpłynąć na udzielane odpowiedzi.

PROSZĘ SIĘ NIE STRESOWAĆ, PYTANIA SĄ PROSTE



Davies i Baker (1987) postanowili więc w dość nietypowy sposób sprawdzić, czy z tym efektem ankietera jest coś na rzeczy – stworzyli

kwestionariusz ankiety, który był skierowany do mężczyzn uzależnionych od heroiny. Pierwszy kwestionariusz został im dostarczony przez powszechnie znanego heroinistę (nie pytajcie mnie o definicję powszechnie znanego heroinisty, najpewniej to taki Robert De Niro świata narkotyków – wszyscy go znali, podziwiali i mówili, że był doskonały w tym *Milczeniu owiec*). Drugi kwestionariusz został im dostarczony 10–14 dni później przez człowieka w ogóle niezwiązanego z tym światem, takiego sobie zwykłego Józka, który rano chodzi po bułki do sklepu, mówi „dzień dobry” sąsiadom i sprząta po swoim psie, a najbardziej szaloną rzeczą, jaką zrobił w życiu, było przejechanie 23 minut na 20-minutowym bilecie (!!!). Badanym mówiono, że oba kwestionariusze nie mają ze sobą nic wspólnego, że są to dwa osobne badania. Odpowiedzi różniły się w zależności od tego, kto przeprowadzał badanie, ale w dość zaskakujący sposób – badani pytani przez pana Józka znacznie częściej opisywali siebie jako bardziej uzależnionych i biorących więcej narkotyków.

Zarówno efekt społecznych oczekiwań, jak i efekt ankietera zdarzają się najczęściej w przypadku, gdy pytamy o tematy trudne lub wrażliwe. Sprawdzili to też McBee i Justice (1976), którzy wysłali ankieterów, by zebrali wywiad psychiatryczny wśród 128 dorosłych pacjentów obojga płci. Tym razem sprawdzano, w jaki sposób zachowanie ankietera wpływa na liczbę zgłaszanych symptomów. Okazało się, że pacjenci byli w tej kwestii najbardziej otwarci w przypadku, gdy ankieter udzielał im wzmacniających, wspierających komunikatów, był ubrany w fartuch laboratoryjny i siedział za biurkiem naprzeciwko badanego. Dajcie spokój, wszyscy wiemy, że jeśli pan w reklamie leku na ból łokcia ma na sobie fartuch lekarski, to wiadomo, że możemy mu ufać. A to się dobrze składa, bo przy okazji możemy sobie do takiego doktora przyjść i poprosić, by reanimował naszą chorą ankietę.

Dej (wskazówki), mam chorom ankiete

Czasem, gdy opowiadam w różnych miejscach o efekcie społecznych oczekiwań lub innych błędach poznawczych, które mogą zrujnować nasze badania ankietowe, spotykam się z reakcją podobną do swojej, gdy mąż pyta mnie, czy chcę gryza kanapki. Z tym że w tej drugiej sytuacji co najmniej jedna strona czuje się potem pokrzywdzona.

Faktycznie, kiedy zapoznamy się z tymi błędami, bardzo często wydają nam się one intuicyjne i logiczne. Niemniej w trakcie tworzenia i analizowania wyników ankiety bardzo często zdarza nam się o tym zapomnieć. Ma to znaczenie również wtedy, gdy czytamy i próbujemy zrozumieć wyniki ankiet przedstawianych w mediach. Na przykład w „Wysokich Obcasach” pojawił się artykuł, według którego 72% Polaków deklaruje, że segreguje śmieci. Ale według danych Polskiego Zakładu Gospodarki Komunalnej (opartych na liczbie tych segregowanych worków, które otrzymują) robi to zaledwie 27%. No i co teraz? Który Czesiek wrzucił ostatnio słoik po jabłkach, co mu matka usmażyła, do odpadów papierowych, bo wymyślił sobie, że wszak były to jabłka, he, he, papierówki?

No słuchajcie, warto się nad tym zastanowić. To zdanie pozostawione samo sobie jest trochę jak taki gumowy łabędź na wzburzonym morzu ludzkich interpretacji⁶² – gdzie zawieje, tam poleci. Jeśli ktoś na podstawie tej jednej informacji będzie chciał uznać, że wszelkie ankiety są bez sensu, a badania naukowe to liche fundament jakiegokolwiek osądu, taki fundament wszelkich wniosków zrobiony z rozmokłej tektury i zardzewiałego widelca, obawiam się, że tak właśnie uzna i wszystkim będzie opowiadał, że badania kłamią, a intuicja, intuicja, proszę pani, to mnie jeszcze nigdy nie zawiodła! INTUICJA I ZDROWY ROZSADEK!!!

Jeśli ktoś na podstawie tego jednego zdania będzie chciał wysnuć wniosek, że Polacy to fleje, nic tylko rzucają starymi ścierkami przez okno, a potem celebują wspólne tarzanie się w obierkach po ziemniakach, to też najpewniej tak właśnie robi. A tak naprawdę to, co jest dla nas wszystkich w tym zdaniu najistotniejsze, to – ośmielę się zasugerować – prawda. Czyli ustalenie, jak jest z tą segregacją. Czemu mamy dwie tak odmienne statystyki? Jak to zinterpretować? Wiemy już, że temu doniesieniu opartemu na deklaratywności możemy ufać tak sobie – spytani o to, czy segregujemy śmieci, czy też codziennie rzucałyśmy foliówkami w żółwie morskie, chętniej odpowiemy, że tak, tak, oczywiście, segregujemy. No bo jednak wolimy w swoim bloku wyjść na czołowych ekologów, żeby potem obciachu nie było, że komuś się nie chce nawet przez chwilę zastanowić nad tym, gdzie wyrzucić papierka, a tymczasem Krystyna spod piątki rzetelnie segreguje odpady, przygarnęła do siebie pięć bezdomnych manatów i postawiła w ogródku ule, bo dowiedziała się, że pszczoły bardzo

potrzebują domów, a żaden hotel nie chce ich przyjąć, bo nie mają kart kredytowych.

Kłopot z tymi ankietami. Pomyślcie o nich jak o kapibarach. Z jednej strony budzą w wielu osobach uczucia ciepłe i pluszowe, no bo jednak wydają się niegroźne i całkiem przystępne, z drugiej strony wciąż na ich temat krąży wiele niedomówień, to znaczy mam tutaj na myśli badania ankietowe, ale w sumie kapibary też, no dajcie spokój, zastanówcie się przez chwilę, co wiecie o kapibarach. Czy wiecie, co jedzą? Co je cieszy? Co smuci? Jaki jest ich ulubiony program w telewizji? Jaki jest ich stosunek do ekonomii politycznej ze szczególnym uwzględnieniem perspektywy pozytywnej?

No i widzicie, teraz wystąpiła pewna niezręczność, bo nagle okazało się, że większość osób niewiele wie o kapibarach. Mam dla was złą wiadomość: najpewniej tak samo jest w przypadku badań ankietowych. Na przykład czy wiecie, że gdyby badania ankietowe zorganizowały wigilię dla całej rodziny, to przybyłoby wielu gości, bo i badanie ankietowe typu CATI, i typu CAWI, i CAPI też przyjdzie, i PAPI też. A każde z nich ma własne kłopoty, o których opowiada rodzinie, łącznie z tym, na ile musi sobie radzić z problemami wynikającymi z deklaracyjności czy efektu ankietera. CATI, CAWI, CAPI, PAPI. Być może myślicie, że to dziwne imiona dla dzieci, ale już wam przedstawiam wszystkich członków tej radosnej rodziny.

CAWI (ang. Computer-Assisted Web Interview)

to wywiad wspomagany komputerowo, czyli że bierzemy sobie takiego miłego człowieka i prosimy go o wypełnienie kwestionariusza on-line, i on go wypełnia samodzielnie, bez ankietera. W tym wypadku nie ma więc osoby, która czyta ankietowanemu pytania, i taki biedny Jędrzej nie musi myśleć, że ktoś go oceni i złoży donos do Świętego Mikołaja, że prezenty się nie należą, bo Jędrzej zadeklarował, że często kopie gołębice na ulicy⁶³.

CATI (ang. Computer-Assisted Telephone Interview)

to wywiad przeprowadzany telefonicznie, w którym ankieter czyta pytania, a potem wklepuje uzyskane odpowiedzi do komputera. To ten przypadek, kiedy ktoś do was dzwoni i mówi: „Halo, czy chciałby Pan porozmawiać

o plastiku?”, a wy się zgadzacie, bo akurat telefon zatrzeszczał i myśleliście, że pytają, czy chcecie porozmawiać o serniku.

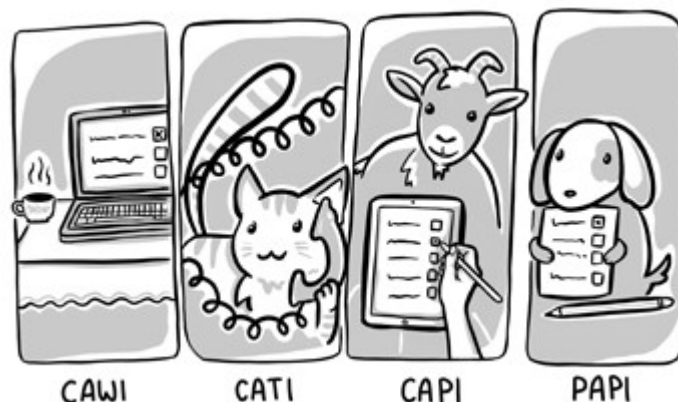
CAPI (ang. Computer-Assisted Personal Interview)

to metoda, w której kwestionariusz wypełnia się znów za pomocą komputera (to jest laptopa, iPada lub innego urządzenia mobilnego, a nie że ankieter targa ze sobą komputer stacjonarny od domu do domu, bo to by było niepoważne, przecież już w okolicach Mokotowa by mu kabla nie starczyło). Niemniej ankieter przychodzi do nas osobiście, czyta pytania, a my mu udzielamy odpowiedzi, które on zaznacza w komputerze. Czyli to w sumie taki kwestionariusz w wersji audiobookowej.

No i nasza ostatnia kwestionariuszowa gwiazda, o nazwie tak pluszowej, że bardziej kojarzy się z imieniem dla puchatego kundelka niż z techniką ankietowania – PAPI.

PAPI (ang. Paper and Pencil Interview)

czyli ten przypadek, kiedy drukujemy kwestionariusze, rozdajemy wśród swoich ankietowanych, a oni je wypełniają ręcznie jak amisy. Uspokajam również, że nazwa tej techniki jest zwyczajowa, to wcale nie jest tak, że jeśli nie macie przy sobie ołówka, jedynie długopis z jednorożcem, który ma pióra zamiast grzywy, to nie wolno wam tego kwestionariusza wypełnić.



I teraz tak, nie opowiadam wam o tych wszystkich technikach i ciekawych skrótach, byście mogli zarapować na ich temat⁶⁴. Piszę o nich,

bo wybór tych technik przypomina trochę wybór potrawy w menu – każda z nich ma określone wady i zalety⁶⁵. I tak efekt społecznych oczekiwań jest najniższy w przypadku technik CAWI i CAPI, albowiem tam nie ma człowieka, który siedzi przed nami i może oceniać, czy nasze odpowiedzi są spoko, czy jednak brzmiemy jak psychopaci. Efekt ten jest też niższy w przypadku, gdy wywiad jest przeprowadzany przez telefon, w porównaniu z sytuacją, w której pytania są zadawane bezpośrednio, twarzą w twarz. No bo jednak kiedy dzieli kogoś od nas całkiem spory dystans, to trochę łatwiej nam się przyznać do rzeczy, z których niekoniecznie jesteśmy dumni. Podobnie jest z efektem ankietera – nie występuje w przypadku kwestionariuszy, które wypełniamy sami (on-line lub na papierze), no bo wiecie... tam nie ma ankietera. Ale efektywność takiego rozwiązania jest mniejsza, bo o ile przy kimś trochę nam głupio przerwać wywiad w połowie i powiedzieć, że ja bardzo przepraszam, ale musimy kończyć, bo właśnie rozpoczął się sezon na szczupaki, no a „szczupak jest król wody, jak lew jest król dżungli”, o tyle gdy jesteśmy sam na sam z ankietą, rzadziej mamy takie dylematy. Zjawisko polegające na tym, że zależnie od tego, jaką technikę wybierzemy, możemy uzyskać różne odpowiedzi (i różną liczbę odpowiedzi), nazywa się *survey mode effect* i zostało wielokrotnie przebadane.

Eksperymenty, które mają sprawdzić, czy technika badawcza w jakikolwiek sposób wpływa na jakość odpowiedzi i „wypełnialność” ankiet, z reguły wyglądają tak samo: dokładnie ten sam kwestionariusz jest rozprawdany wśród dwóch lub większej liczby grup, ale różnymi technikami. To znaczy, że badani są proszeni o wypełnienie kwestionariusza przez internet, na papierze lub jeszcze inaczej, a potem sprawdza się, czy w danych grupach wystąpiły różnice w „wypełnialności” i jakości udzielanych odpowiedzi.

W jednym z takich badań porównywano kwestionariusz wypełniany internetowo (CAWI) z takim, w którym pytania były zadawane osobiście przez ankietera (CAPI). Kwestionariusz internetowy wypełniło 5,1% respondentów (czyli 192 na 3785 wysłanych zaproszeń). Kwestionariusz, w którym pytania były czytane przez ankietera, wypełniło 84% ankietowanych, czyli 230 z 275 osób (Marta-Pedroso et al. 2007). Inne badanie (Clariana i Wallace 2002), wykazało, że grupa wypełniająca kwestionariusz w wersji komputerowej poradziła sobie zdecydowanie lepiej niż grupa, która wypełniała identyczny kwestionariusz w wersji

papierowej, a różnica ta nie zależała od płci respondenta, jego ambicji czy umiejętności obsługi komputera.

Trzeba również pamiętać, że dla „wypełnialności” ankiety ma znaczenie nie tylko technika, której używamy, czy ankieter, którego wysyłamy w świat do zadawania pytań. Błędów poznawczych związanych z badaniami ankietowymi jest znacznie więcej⁶⁶. Udowodniono, że na to, jak odpowiadamy, wpływają między innymi: kolejność pytań (*question order effects*)⁶⁷; to, o jak odległe w czasie wydarzenia pytamy i jak bardzo nadwyrężamy pamięć respondenta (*recall bias*); to, jak wiele pytań zadajemy i jak bardzo są one wymagające dla ankietowanego (*response fatigue*); oraz to, w jaki sposób formułujemy pytania (*question wording bias*)⁶⁸. Moim ulubionym błędem poznawczym jest zaś *haptic metaphor effect*, który występuje nie tylko w badaniach ankietowych, lecz także jest tak szalenie interesującym błędem, że bezapelacyjnie zasługuje na przytoczenie w tej książce. Występuje on mianowicie wtedy, kiedy na nasze odpowiedzi wpływają... dotyk i wszelkie fizyczne odczucia. Na przykład jeśli chwilę przed spotkaniem z inną osobą potrzymany w dłoniach gorący napój, to częściej ocenimy naszego towarzysza jako osobę o ciepłym usposobieniu, niż gdybyśmy sobie przed tym spotkaniem potrzyмали kostki lodu⁶⁹.

Jesteśmy tylko ludźmi, i tak nie unikniemy wielu błędów, które mogą nam się przytrafić w czasie przeprowadzania badania lub samego wypełniania kwestionariuszy. Niemniej daleka byłabym od rezygnacji z tego powodu z badań ankietowych. W końcu gdybyśmy to zrobili, to nigdy byśmy się nie dowiedzieli, że 11% (z 2392 ankietowanych) Amerykanów myśli, że HTML to choroba przenoszona drogą płciową⁷⁰.

Poza tym są też dobre wiadomości – metodologicznie rzecz ujmując, trudności wynikające z zadawania wrażliwych pytań, na przykład tych o seks, dałoby się łatwo skorygować. I nie, nie mam tu na myśli obserwacji uczestniczącej w celu upewnienia się, że ankietowani odpowiadają zgodnie z rzeczywistością. Warto się nad efektem społecznych oczekiwań, efektem ankietera i wszelkimi innymi błędami związanymi z badaniami ankietowymi zastanowić jeszcze na etapie planowania badania. Wtedy też możemy zdecydować, czy rzeczywiście kwestionariusz ankiety jest najlepszym sposobem na dowiedzenie się, co ludziom w głowach, sercach i uprzedzeniach siedzi. Tam, gdzie jest to możliwe, problem

deklaratywności można rozwiązać na przykład dzięki zastosowaniu metody eksperymentalnej. W takim przypadku najpierw trzeba znaleźć odpowiedź na jedno szalenie ważne pytanie, mianowicie: czy na imieniny lepiej kupić czekoladki, czy kokainę?

Na imieniny lepiej kupić czekoladki czy kokainę?

Wymyśliłam sobie, że znowu napiszę coś o swojej robocie na uniwersytecie, bo przecież jak ja nie napiszę o tej wspaniałej intelektualnej przygodzie olinkluziw i lastminyt w jednym, o tym najdelikatniejszym z biszkoptów pedagogicznej wuzetki, najpuchatszej owcy na edukacyjnym pastwisku wiecznej szczęśliwości, to nikt nie napisze.

To znaczy moim studentom co prawda nie brak talentu, ale oni raczej z delikatnego kruszcu słów rzeźbią historie barokowo piękne i zawile o tym, jak nie mogli dotrzeć na zajęcia, bo w drodze na uczelnię zatrzęsała się ziemia, pękać jak krucha wydmuszka zaczęła jej skorupa, na ich oczach rozstępować się zaczęły płyty tektoniczne, utworzywszy następnie ogromny wąwóz wypełniony ogniem i śmiercionośną lawą, a nad tym wąwozem unosiły się kawałki skał, samochodów i szczątków tych, którzy próbowali przedostać się na drugą stronę. Ale to ich oczywiście nie powstrzymało, oni nawet na chwilę nie ustali w swoich heroicznym próbach dotarcia na uczelnię i przedzierali się przez tę lawę, ten wąwóz ognia, ten pył ofiar poległych w starciu z dzikością natury, przez te krzewy kolczaste złych zbiegów okoliczności, a pot mieszał im się z krwią, a krew ze łzami – artefaktami rozpacz, że już umknęło im pierwsze pięć minut moich zajęć. Parli w kierunku uczelni, wbrew rozsądkowi i wbrew wszelkim żywiołom, no ale potem się zgubili, bo źle skręcili w objazd. I oni mi piszą takie właśnie maile, i potem jeszcze pytają, czy z okazji tych gejzerów ognia i lawy mogłabym im usprawiedliwić absencję na moich wyśmienitych zajęciach, a ja im na to zawsze odpisuję to samo, to znaczy, że ja przecież na tym przedmiocie od trzech lat nie sprawdzam obecności.

Wiecie, istnieją takie miesiące, kiedy to w robocie spędzam bardzo dużo czasu na słuchaniu Celine Dion i tego utworu, który serce na kawałki drze jak kot bibułę, czyli *My heart will go on*. Ta piosenka doskonale współgra z fałszywymi nutami mego pochmurnego życia, wplata się ścięciem dokładnym w moją wyszywaną codziennej rozpacz, łyżką goryczy miesza w roztworze z cytryny mojej egzystencji. Chodzi mi oczywiście o czas egzaminów.

Podczas egzaminów zawsze jestem jak Leonardo DiCaprio dryfujący w oceanie błędnie wyliczonych odchylen standardowych, oceanie lodowatym i paraliżującym tak, że aż człowiek zapomina, że aby przeżyć,

wystarczy przecież wejść na tę krę. A musicie wiedzieć, że jak się człowiek już raz w takim oceanie lodu i braku nadziei wykąpie, to wcale nie chce do niego wracać w czasie sesji poprawkowej i tutaj właśnie przydać się miała jedna z wielu moich zalet, która polega na tym, że ja niezmiennie mam tak doskonałe pomysły, że aż czasem podejrzewam, że jestem nieślubnym dzieckiem Napoleona i projektanta mebli IKEA.

I właśnie dlatego swego czasu opracowałam przemyślany system egzaminacyjny, który składał się z trzech kroków. Po pierwsze, podałam moim studentom trzy pytania, które się na tym egzaminie pojawią. Następnie podałam im odpowiedzi. A potem jeszcze – w przypływie geniuszu strategicznego – na egzaminie kazałam im z tych trzech pytań wybrać sobie tylko jedno, posypałam to cukrową posypką progu zdawalności 40% i to był tak wyborny plan, to była tak przemyślana strategia, że na efekty nie trzeba było długo czekać: oblało pięć osób.

Ja nie wiem, jak to się stało, nie pytajcie mnie. To by zdał każdy gołąb, nawet gdyby go w trakcie egzaminu rozkojarzyć kromką chleba.



Niemniej w obliczu tak dramatycznej ewaluacji moich osiągnięć pedagogicznych postanowiłam zrobić to, co każdy zrobiłby na moim miejscu – wyjść z uczelni i już nigdy nie wrócić. No to idę ulicą. Moje życie jest jak budyń czekoladowy z obniżoną zawartością cukru, czyli bez sensu. Jestem najsmutniejszym człowiekiem na tym dansingu rzeczywistości, tym brzydkim chłopcem w zamszowym krawacie, co na ten dansing przyszedł bez dziewczyny, ale za to w towarzystwie delikatnego zapachu środka na mole i ludzkiej pogardy, a potem przez godzinę stał w kolejce do bufetu i jak już tam w końcu dotarł, to ktoś mu ukradł ostatniego koreczka ze śledziem i serem. Gorzej być nie może – tak właśnie sobie myślę, że gorzej

być nie może, ale to nigdy nie jest prawda. Gorzej być nie może – powiedział Zygmunt spod Pcimia, jak mu w 1772 roku krowa zdechła. A potem przyszły rozbiory.

Idę i jest mi smutno, a kątem oka widzę panią ankieterkę, widzę, że czai się na mnie jak wąż w pomidorach, ale nie mam siły walczyć, proszę pytać, proszę pani, proszę pytać, o co tylko pani chce, ja na wszystkie pytania chętnie odpowiem. Jak żyć? Jaki jest sens zła? Dokąd zmierzamy? Już jestem gotowa odpowiedzieć na wszystkie te pytania, ale jest lepiej, bo kiedy mnie w końcu dopada, to mówi, że ona ma do mnie tylko jedno pytanie, bardzo krótkie pytanie, i ona wcale nie chce dyskutować o teodycei ani przedplatońskiej koncepcji dobra, ona tylko jedno pytanie do mnie ma... – tu moduluje głos, patrzy na mnie przeciągle, ołówek wymownie w powietrzu zawiesza i w końcu pyta: Czy myślała pani kiedyś o operacji plastycznej?

No proszę pani, przecież jest wiele miłszych sposobów, by powiedzieć, że nie wyglądam dziś najlepiej. I wiele lepszych metod na to, by zebrać dane dotyczące tematów, o których ludzie nie lubią mówić.

Panie doktorze, czy to HTML?

Spójrzmy na przykład na wspomniane już wcześniej i głośne swego czasu badanie, które wykazało, że co dziewiąty Amerykanin myśli, że HTML to choroba weneryczna, he, he, he. To zawsze radość, gdy okazuje się, że 11% populacji⁷¹ to ludzie głupszy od nas, nawet jeśli w ostatniej krzyżówce, którą rozwiązywaliśmy, w odpowiedzi na hasło „graniczy z Polską” wpisaliśmy „Sosnowiec”. Taki to był chwytliwy temat, że aż omówiły go CNN, „Independent” oraz „Time” (choć tylko ten ostatni pokusił się o jakąkolwiek refleksję dotyczącą badania). Niewielu pomyślało o tym, skąd taki wynik się wziął i czy można mu ufać. Kwestionariusz został przeprowadzony wśród czytelników strony oferującej kupony zniżkowe i miał sprawdzić znajomość terminów technologicznych – dokładniej rzecz ujmując, takich związanych z tworzeniem stron WWW. Mimo że badani mieli do wyboru trzy odpowiedzi, to i tak poszło im – jak donosiły media – zenująco. 11% myślało, że HTML to choroba weneryczna, 18% – że Blu-ray to morskie zwierzę, a 27% – że gigabajt to owad, którego można spotkać w Ameryce Południowej.

Po pierwsze, próba została pobrana w sposób celowy, a więc badanie nie było reprezentatywne. Możliwe, że grupa badana różniła się w jakiś sposób cechami demograficznymi od reszty populacji⁷² (na przykład ze stron z voucherami i kuponami częściej korzystają osoby starsze). Po drugie, dobrze byłoby wiedzieć, w jaki sposób zostały zadane pytania, jak dokładnie brzmiały (tej informacji nigdzie nie udało mi się znaleźć). To zawsze ma ogromne znaczenie – nie tylko wtedy, gdy pytania w jakiś sposób sugerują respondentowi odpowiedź⁷³. Jak widzicie, przed przeprowadzeniem ankiety badacz powinien wziąć pod uwagę bardzo wiele czynników, tak samo jak my, gdy weryfikujemy, na ile możemy ufać jej wynikom. Sprawa komplikuje się jeszcze bardziej, gdy przychodzi nam pytać o rzeczy dla ludzi niewygodne.

Pytania na tematy wrażliwe, społecznie nieakceptowane i objęte tabu to zagadnienie od zawsze metodologicznie trudne. Można by powiedzieć – takie, które od lat spędza sen z powiek metodologom (znaczy, gdyby oni w ogóle spali, a nie byli ciągle zajęci rozmyślaniami o tym, czy skala w pytaniu ankietowym powinna być pionowa; czy pozioma⁷⁴). Nie tylko dlatego, że ludzie kłamią celowo, jak w przypadku efektu społecznych oczekiwań czy efektu ankietera. Czasem po prostu trudno im sobie przypomnieć określone zdarzenia z przeszłości lub wspomnienia o nich są zniekształcone⁷⁵. Inną sprawą jest to, że wielu z nas może mieć nieświadomione uprzedzenia (ang. *unconscious bias*), czyli mimowolne negatywne nastawienie do danej grupy czy osoby lub też mimowolną przychylność wobec innej. Wszystkie te problemy wynikające z deklaratywności, nieświadomych uprzedzeń, efektu społecznych oczekiwań czy trudności związanych z pytaniami wrażliwymi wymagają szczególnego podejścia metodologicznego.

Wiemy już, że nie warto podchodzić do ludzi i pytać: „Dzień dobry, czy lubi Pan/Pani dzieci?”. Zwłaszcza jeśli przypadkiem obok przechodzi grupa z okolicznego przedszkola⁷⁶. Jednym ze sposobów poradzenia sobie z problemami, o które trudno spytać w ankiecie, jest zastosowanie metod eksperymentalnych lub quasi-eksperymentalnych. Bo musicie wiedzieć, że ta metoda badawcza jest zarezerwowana nie tylko dla siwych panów z rozwianymi włosami, którzy biegają po laboratorium w białych fartuchach, wyśpiewując ciąg Fibonacciego – równie dobrze spełnia się w naukach społecznych i psychologicznych. Prawdą jest natomiast, że

czepie ona garściami z tego, co wypracowały nauki ścisłe i medyczne. Na przykład z metodologii badania, które miało odpowiedzieć na kluczowe dla wszystkich pytanie, czyli: co najlepiej przynieść w prezencie szczurowi Bartkowi, gdy zostaniemy przez niego zaproszeni na rodzinną uroczystość?

Co najbardziej jara szczura Bartka?

Nie no, bądźmy poważni. Wyobraźmy sobie taką sytuację: dzwoni do nas żona szczura Bartka i zaprasza na jego imieniny, a zapytana o wskazówki dotyczące prezentu odpowiada, że wystarczy jakiś drobiazg, coś, co wrzucić można na – nomen omen całkiem spory – ząb. Co powinniśmy sprawić solenizantowi? Czy lepiej iść w klasyczną bombonierkę, czy też bardziej ucieszy go, no nie wiem, kokaina? Słuszne pytanie! Ale nie tylko my mamy ten problem. W 2017 roku zastanawiali się też nad nim francuscy, badacze Lenoir i spółka. Co więcej, przeprowadzili eksperyment, który miał dać odpowiedź na to pytanie, a wyniki opublikowali w artykule *Intense Sweetness Surpasses Cocaine Reward*⁷⁷.

By to sprawdzić, przeprowadzili następujący eksperyment: losowo podzielili szczury na trzy grupy, a następnie reprezentantom każdej z nich dali do wyboru dwie dźwignie. W tym eksperymencie było jak w życiu – niektórym to brak kolejki w mięsnym i zawsze autobus przyjeżdża na czas, a innym wiecznie pod górkę w za ciasnym berecie. W pierwszej grupie szczury po naciśnięciu jednej z dźwigni były nagradzane sacharyną (to jest słodzikiem, kilkaset razy słodszy od cukru), a po naciśnięciu drugiej nie były nagradzane wcale. W drugiej grupie słodka nagroda została zastąpiona dawką kokainy (pozostała dźwignia wciąż oferowała jedynie pustkę, marazm i rozczarowanie). No i w końcu trzecia grupa szczurów miała wybór, przed którym raczej nie staniemy w spożywczaaku – jedna dźwignia oferowała im słodzik, a druga narkotyk. Opisuję tutaj uproszczoną wersję eksperymentu⁷⁸, niemniej wystarczającą dla naszych rozważań. Wyniki były wyjątkowo korzystne dla słodkiego smaku – gryzonie wolały cukier od kokainy (bez względu na to, czy był to sztuczny słodzik, czy też naturalny cukier⁷⁹). Nawet szczury gustujące w kokainie (te uzależnione) dość chętnie przerzucały się potem na cukier. Co więcej, były w stanie zapracować na niego znacznie ciężiej niż na narkotyk.

Co mnie w sumie nie dziwi, bo ja też przez wiele lat swojej rozkosznie pedagogicznej drogi testowałam efekt nagradzania studentów białym

proszkiem w zamian za poprawne odpowiedzi. Chodzi mi oczywiście o cukier. Wyniki moich prób były mieszane, to znaczy okazało się, że w obliczu cukierków studenci owszem, stają się bardziej aktywni i chętniej się udzielają w czasie zajęć, niemniej z czasem ich uwaga rozprasza się wskutek licznych prób wypowiedzenia słowa Michaszki.



Badanie eksperymentalne i wszyscy jego liczni kuzyni

Badanie na kochających cukier i kokainę szczurach jest klasycznym schematem eksperymentalnym, który charakteryzuje się następującymi cechami:

- **OBECNOŚĆ CO NAJMNIEJ DWÓCH GRUP BADAWCZYCH, Z CZEGO JEDNA JEST GRUPĄ KONTROLNĄ.** Grupa kontrolna to taka, w której nie manipulowano żadną zmienną, na przykład w przypadku wyżej opisanego badania – obecnością cukru lub kokainy. To znaczy, że w grupie kontrolnej nie stosujemy żadnych oddziaływań, nie wprowadzamy żadnych zmian w porównaniu z grupami eksperymentalnymi, w których manipulujemy interesującym nas czynnikiem (na przykład obecnością lub dawką cukru).

- **BADANI (nieważne, czy są to ludzie, szczury, miasto, czy jakiegokolwiek inne obiekty badane) SĄ PRZYDZIELANI DO GRUP W SPOSÓB LOSOWY (*random assignment*)⁸⁰.** To oznacza, że każdy z badanych może się znaleźć w każdej z grup z takim samym prawdopodobieństwem (decyduje o tym na przykład wynik losowania czy rzut monetą, a nie sam badacz).

- KONTROLA NAD ZMIENNYMI NIEZALEŻNYMI (w tym przypadku słodzikiem i kokainą). W wyżej opisanym badaniu badacz miał kontrolę nad tym, czy dany szczur otrzyma daną substancję, czy nie, a w dalszych etapach badania również nad jej dawką.

W naukach ścisłych mamy czasem do czynienia z eksperymentem również wtedy, gdy nie tworzymy grup eksperymentalnych i kontrolnej, ale zachowany jest warunek pełnej kontroli nad środowiskiem badania. Przykładem może być najdłuższy eksperyment świata, czyli eksperyment kropli paku. Eksperyment ma udowodnić, że w warunkach standardowych (STP – *Standard Temperature and Pressure*) pak⁸¹ nie jest ciałem stałym, lecz cieczą o tak ogromnej lepkości, że formuje ona jedną kroplę przez wiele lat (dokładniej rzecz ujmując – około dziesięciu). Eksperyment jest prowadzony na Uniwersytecie Queensland w Brisbane w Australii od 1927 roku (choć z przerwami i wprowadzonymi po drodze niewielkimi zmianami). Pak został wlany do lejka, skąd może swobodnie wypływać. Liczy się to, jak długo trwa uformowanie kolejnych kropli.



Thomas Parnell, twórca eksperymentu, zmarł w 1948 roku i dożył upadku dwóch kropli⁸². Ostatnia, już dziewiąta, kropla spadła w 2014 roku. Podobny eksperyment jest prowadzony na mojej macierzystej uczelni, Trinity College w Dublinie, co czyni mnie szalenie dumną, bo to oznacza, że mój doktorat nie jest najdziwniejszym badaniem, które tam zaprojektowano. Trwa od 1944 roku (w sensie eksperyment, nie mój doktorat) i co ciekawe – pierwsza kropla spadła dopiero w 2013 roku. Oba eksperymenty – i ten w Australii, i ten w Irlandii – są obecnie stale

monitorowane przez szereg kamer⁸³, bo okazało się, że na posadę obserwatora formujących się przez kilkanaście lat kropli paku nie ma zbyt wielu chętnych (najpewniej dlatego, że nie oferowano wejściówek do muzeum kolejnictwa i kuponów na cheeseburgery).

I jeszcze jedna rzecz. W badaniach eksperymentalnych przeprowadzanych – nie tylko, lecz głównie – w medycynie czy w farmacji szalenie istotne jest maskowanie badania, czyli pojedyncza, podwójna lub potrójna ślepa próba. W tej pojedynczej jedynie pacjent nie wie, czy otrzymuje leczenie, czy placebo, w podwójnej – zarówno pacjent, jak i lekarz nie mają tej informacji, a w potrójnej – nie posiadają jej ani pacjent, ani lekarz, ani laborant (lub statystyk) analizujący wyniki badania. Cel to zwiększenie dokładności i obiektywności badania oraz zminimalizowanie wszelkich zakłóceń wynikających z uświadomionych i nieświadomych działań osób zaangażowanych w badanie.

Niemniej pozostałe warunki badania eksperymentalnego są jak najbardziej spełnione – badacze świadomie manipulują czynnikami lub bodźcami, by sprawdzić, czy zastosowanie takiej manipulacji wpłynie na zmianę zachowania lub stan zdrowia badanych. Nie decydują jednak o przydziale badanych do grup. A jak ta historia o czekoladzie i narkotykach ma się do naszych początkowych rozważań o zbieraniu danych na tematy trudne i wrażliwe? No słuchajcie, wbrew pozorom ma z nimi bardzo wiele wspólnego – często to właśnie zastosowanie metod eksperymentalnych w badaniach społecznych pozwala nam znaleźć rzetelną i wiarygodną odpowiedź na pytania, na które uczestnicy nie bardzo lubią odpowiadać. Na przykład takie, na które znajdziemy odpowiedź w kolejnym rozdziale, będącym kontynuacją rozważań o tym, jak badać tematy wrażliwe, czyli: czy wszyscy jesteśmy rasistami?

Czy wszyscy jesteśmy rasistami?

Problem z trudnymi pytaniami polega na tym, że wrażliwość danego tematu oraz to, czy jest uznawany za tabu, zależą zarówno od osobistych uwarunkowań rozmówcy, jak i kontekstu kulturowego (różnica w obrębie poszczególnych nacji etc.). Zresztą dajcie spokój, nie trzeba szukać wśród tematów trudnych, poszczególne narody nie zgadzają się między sobą nawet w kwestii ustalenia definicji ładnej pogody. Mówię to jako człowiek, który prawie dekadę spędził w innym kraju – gdzie nauczyłam się wiele nie tylko o różnicach kulturowych, lecz także o meteorologii.

Bo gdybym miała wymienić jedną rzecz, za którą najbardziej cenię Irlandię, to bez wątplenia byłaby to pogoda. Zwłaszcza latem, bo wtedy deszcz staje się cieplejszy, a człowiek może nosić kurtkę przeciwdeszczową i sandały jednocześnie. Niemniej od czasu do czasu przydarzają się w Irlandii takie trzy dni w roku, kiedy przestaje padać, na niebie pojawia się słońce, na termometrze 15 stopni (!!!) i tym samym na pierwszej stronie wszystkich gazet można przeczytać: *HEAT WAVE*. Często korzystałam z tej pięknej, słonecznej aury i zabierałam swoich studentów na zajęcia na dworze, głównie dlatego, że jak czasem patrzyłam na to, co liczą, to miałam przecucie graniczące z pewnością, że cierpią na jakieś zaawansowane niedobory tlenu⁸⁴. Poza tym zajęcia na dworze są zawsze spoko, bo zaczynasz je z piętnastoosobową grupą studentów, a kończysz z osiemnastoosobową, składającą się ze studentów, dwójki niemieckich turystów i jednego bezdomnego, z czego z reguły bezdomny wykazuje większe zainteresowanie zajęciami niż niemieccy turyści, o studentach nie wspominając. No ale też rzadko się zdarza, by moi studenci po lekcjach prosili mnie o euracza na piwo.

No więc kiedyś wysłałam moim milusińskim maila, że na następnych zajęciach nie będziemy nic liczyć, więc wszyscy ci, którzy wcześniej poinformowali mnie, że są obłożnie chorzy i nie dotrą na zajęcia, mogą bezpiecznie wyzdrowieć. Że będziemy rozmawiać o egzaminie, więc mogą mi przesłać pytania, na które pragną poznać odpowiedź, z tym że wolałabym, żeby były to pytania trochę bardziej sensowne niż te, które zadawali mi w ostatniej anonimowej ankiecie, kiedy to jeden z nich chciał się dowiedzieć, czy kokosa można przesłać pocztą, a drugi – jaki jest sens

życia. I że już sprawdziłam, kokosa można przesłać pocztą, a sensem życia jest jedzenie (niekoniecznie kokosów).

Następnie ustaliliśmy podstawowe zasady naszej współpracy: po pierwsze, ja zobowiązuję się do tego, że nie będę im kazała nic liczyć, a oni, że gdyby tak akurat pobliską alejką przechadzał się mój szef, to wtedy natychmiast złączą słuchać mnie w zachwycie i nie miałabym też nic przeciwko temu, gdyby komuś pocięła łza wzruszenia nad pięknem naszego intelektualnego spotkania. Po drugie, oni powstrzymają się od komentarzy w stylu: *I want to die* oraz wznoszenia dramatycznych okrzyków do Boga typu: *Why do you hate me?!*, w zamian za co ja zobowiązuję się powstrzymać od komentarzy w stylu: *I want to die* oraz wznoszenia dramatycznych okrzyków typu: *Why do you hate me?!*

I wszyscy się zgadzają, wszyscy są zadowoleni, więc zaczynamy. Ja ich pytam, czy pamiętają, do czego służy współczynnik korelacji rang Spearmana, a wtedy oni natychmiast zaczynają patrzeć w prawo, w lewo, w górę, w dół, wszędzie, tylko nie na mnie, więc wyłamiam jakiegoś ochotnika bohatera i pytam człowieka spod krzaka, czy pamięta, do czego służy współczynnik korelacji rang Spearmana, a on na to mówi, że pozwoli sobie odpowiedzieć na moje pytanie pytaniem: do czego mu się to w ogóle przyda?!

A ja mówię, że przyda mu się to na przykład wtedy, gdy przyjdzie do niego jego pięcioletnie dziecko i spyta go: „Tato, tato, a co to jest korelacja rangowa?”, i nie będzie musiał się przed tym dzieckiem wstydzić, a on mi na to, że mój argument jest bez sensu, bo on nie będzie mieć dzieci, bo jest gejem, a ja mówię, że jego argument jest bez sensu, bo równie dobrze może sobie adoptować dziecko, przed którym będzie się wstydzić, a on mi mówi, że mój argument jest bez sensu, bo on w tym kraju nie może adoptować dziecka, a wtedy koleżanka obok mówi, że mogłby jechać do Belfastu i tam adoptować, w końcu pociągiem to tylko trzy godziny, a wtedy inny kolega mówi, że mogłby też jechać do tego Belfastu autobusem, bo to znacznie wygodniej i mają wi-fi, a kolega obok się całkowicie zgadza i jeszcze mu radzi, że powinien też kupić bilet przez internet, to wtedy oszczędzi do kilkunastu euro, a wtedy koleżanka mówi, że w takim razie chyba serio najlepiej będzie, jak on pojedzie autobusem, a nie pociągiem, a ja mówię, że moglibyśmy też na przykład wrócić do korelacji rangowej, ale nikt mnie nie słucha, no bo mówię nie na temat. Wyjątkiem był mój ulubiony student Rudy, który dokładnie w tym momencie krzyknął: *A CAT!*. Wiadomo.

Więc ja mu mówię, że Boże mój, jaki znów kot, a on mówi, że właśnie zobaczył w krzakach kota, a ja go pytam, skąd w nim taka ekscytacja, co on nigdy w życiu kota nie widział, a on mówi, że co prawda widział, ale akurat koty są szalenie ekscytujące, zupełnie odwrotnie niż korelacja. I mnie pyta, czy może iść poszukać tego kota, a ja mówię, że jest mi już absolutnie wszystko jedno, niech idzie szukać tego kota, a wtedy kolega spod krzaka pyta, czy on może iść mu pomóc szukać kota, a ja mówię, że dobrze, niech idą szukać tego kota, i oni idą.

W tym czasie kolega od dziecka już się zastanowił i mówi, że on jednak nie chce adoptować dziecka z Wielkiej Brytanii, bo oni tam mówią jak idioci, nie da się ich zrozumieć, a wtedy koleżanka obok mówi, że w takim razie mógłby sobie adoptować dziecko ze Sri Lanki i tym samym być jak Angelina Jolie, a on mówi, że to w sumie spoko pomysł, ale chyba najpierw musiałyby się dowiedzieć, gdzie leży Sri Lanka, a kolega obok mówi, że osobiście to on by się bał, że z takiego azjatyckiego dziecka to inne dzieci będą się w przedszkolu śmiały, no bo na przykład z rudych dzieci się często śmieją, a ja w tym momencie patrzę na mojego rudego studenta, który biegnie przez boisko do rugby w pogoni za kotem, i myślę sobie, że akurat jeśli chodzi o rudych, to sami sobie pracują na własny PR.

Ogólnie rzecz ujmując, byłam szalenie zadowolona z przebiegu zajęć, myślę, że zrealizowaliśmy bardzo wiele istotnych celów kształcenia. Kolega spod krzaka – gdzie leży Sri Lanka. Jego koleżanka – że przez ostatnich pięć lat przepłacała za bilety do Belfastu. Rudy – że koty bardzo szybko biegają⁸⁵. No i ja – po pierwsze, że istnieje powód, dla którego w tym kraju ciągle pada – by oszczędzić wielu nauczycielom traumy prowadzenia zajęć na dworze.

Po drugie, że gdzieś tam istnieją kraje, w których tematy homoseksualizmu, związków partnerskich i adopcji są traktowane tak naturalnie jak rozmowy o pogodzie. Łatwo jednak wymienić również takie miejsca na Ziemi (a nawet w Europie, wystarczy wyjrzeć za okno), w których ten temat wiąże się z dużo większym tabu i (niestety) ostracyzmem, a tym samym obarczony byłby znacznie większym prawdopodobieństwem efektu społecznych oczekiwań lub innymi błędami, które wpływają na rzetelność wyników badania i ich zgodność z rzeczywistością. Podobnie jest z innymi uprzedzeniami – tymi na tle płci, rasy, religii lub jakimikolwiek innymi. Szczęśliwie na te problemy metodologiczne nauka ma rozwiązanie. Przedstawi nam je teraz Jamal,

który szukał pracy na całkiem rzeczywistym amerykańskim rynku pracy. Czyli wszystko normalnie, może poza tym, że Jamala nie było.

O lepszych i gorszych imionach (znowu). No i o rażeniu prądem (też znowu).

W 2003 roku Marianne Bertrand i Sendhil Mullainathan zastanawiali się nad dokładnie takim samym zagadnieniem jak my, czyli jak sprawdzić, czy na amerykańskim rynku pracy wciąż panuje dyskryminacja na tle rasowym. Wiedzieli, że nie mogą sobie ot tak pochodzić po pracodawcach i spytać. Po pierwsze, to by było bardzo dużo chodzenia, a po drugie, liczyli się z tym, że na pytanie postawione wprost większość osób odpowie przecząco, bez względu na swoje prawdziwe poglądy. Postanowili więc przeprowadzić to badanie inaczej – rozpoczęli od stworzenia fikcyjnych CV. To znaczy całkowicie fikcyjnych, a nie takich trochę pościemnianych, jak wtedy, gdy ktoś aplikuje na stanowisko nauczyciela akademickiego i wpisuje, że przez pięć lat wykładał chemię, co by się nawet zgadzało, z tym że tę chemię to on wykładał na półki w supermarkecie, a nie na uniwersytecie⁸⁶. CV były fikcyjne, ale stworzone w bardzo realistyczny sposób, i stanowiły odpowiedź na prawdziwe ogłoszenia zamieszczone na portalach internetowych⁸⁷. Na każde ogłoszenie odpowiedziano czterema CV – dwoma wysokiej jakości i dwoma niskiej. Różniły się one jedną kluczową zmienną – imieniem aplikanta. Połowa imion sugerowała typowo afroamerykańskie pochodzenie, a połowę stanowiły imiona nadawane częściej ludziom o białym kolorze skóry⁸⁸. Jako że pozostałe czynniki kluczowe dla znalezienia pracy (typu wykształcenie i edukacja) były do siebie bardzo podobne, to łatwo domyślić się, jaki był cel tego badania – sprawdzić, czy pochodzenie (którego pracodawcy domyślali się na podstawie takiego a nie innego imienia) ma wpływ na liczbę rozmów kwalifikacyjnych. Okazało się, że niestety ma – osoby z imionami, które brzmiały jak należące do Afroamerykanów, otrzymały o 50% mniej zaproszeń na rozmowy⁸⁹. Bez względu na jakość CV⁹⁰!

Smutny to wniosek, ale i najpewniej zupełnie inny od tego, co zostałyby zadeklarowane przez tych samych badanych. Tak jak w przypadku innych badań eksperymentalnych, tutaj również mamy do czynienia z grupami eksperymentalną i kontrolną oraz z losowym przydziałem do grup. Różnica

polega na tym, że w tym wypadku badanie jest realizowane w naturalnym środowisku badanych – w ich miejscu zamieszkania, pracy albo nauki. Taki rodzaj badania nazywamy EKSPERYMENTEM TERENOWYM. Jest to technika bardzo często używana w celu zbadania uświadomionych i nieuświadomionych uprzedzeń – nie tylko rasowych, lecz także tych dotyczących wieku (zobacz na przykład Drydakis et al. 2018, Riach i Rich 2006 czy Riach i Rich 2007), płci (Albert i Fernandez-Cornejo 2011, Booth i Leigh 2010, Petit 2007), orientacji seksualnej (Drydakis 2009) czy też pobytu w poprawczaku (Baert i Verhofstadt 2014). Zresztą przeprowadzanie takich eksperymentów nie ogranicza się jedynie do badania rynku pracy. Eksperymentem terenowym jest również jeden z najśłynniejszych eksperymentów w historii psychologii. Nie, nie stanfordzki eksperyment więzienny (zwany także eksperymentem Zimbardo)⁹¹, w którym wsadzano ludzi do wyimaginowanych więzień⁹². Chodzi o ten, w którym ludzie razili się wzajemnie prądem.

To jest psychologia, tego nie ogarniesz.

Stanley Milgram wymyślił sobie, że chce zbadać posłuszeństwo ludzi wobec autorytetu⁹³. Gdyby po prostu o to spytał, to wiadomo, że wszyscy odpowiedzialiby, że tak, wiadomo, autorytet rzecz ważna. Oczywiście oprócz dwulatków, które powiedziałyby, że nie, no bo one na wszystko mówią „nie”.

Zaprojektował więc eksperyment, który następnie powtórzył w kilku lokalizacjach. Zbadał łącznie 1000 osób o różnych cechach demograficznych (takich jak wiek, wykształcenie czy płeć). Pierwszy wariant badania przeprowadził na studenckich ochotnikach bohaterach, którzy zgłosili się po przeczytaniu ogłoszenia w prasie. Prawdą jest, że za udział dostali ekwiwalent około 30 dzisiejszych dolarów. Entuzjazm ochotników wcale nas zresztą nie dziwi, biorąc pod uwagę, że każdy miał kiedyś jakiegoś kolegę Karola z akademika, który to kolega za dwa złote zgodził się wypić trzy litry smalcu i zagryźć kawałkiem balustrady z pobliskich schodów. Niemniej na usprawiedliwienie tych miłych studentów powiem, że wtedy jeszcze nie wiedzieli, że będą razić ludzi prądem, powiedziano im raczej, że będą brać udział w badaniu wpływu kar na pamięć⁹⁴. Wyglądało ono tak, że osoba badana wchodziła do pomieszczenia, z którego miała widok na drugą osobę – w rzeczywistości pomocnika badacza. Podczas losowania ról wszystkim osobom badanym przypadała rola „nauczyciela”, najpewniej dlatego, że losowanie było

ustawione. Ich zadanie polegało na czytaniu drugiej osobie par słów, a potem sprawdzaniu, jak wiele z nich zostało zapamiętanych przez „ucznia”. W przypadku poprawnej odpowiedzi „nauczyciel” czytał kolejną parę, w przypadku złej – dostawał polecenie, by porazić „ucznia” prądem, a z każdą kolejną pomyłką zwiększać siłę wstrząsu. Badani mieli do dyspozycji 30 przełączników, pod którymi umieszczone były opisy siły wstrząsu – od słabego, przez bardzo silny, aż po poważny wstrząs, stanowiący niebezpieczeństwo. W chwili, gdy „nauczyciele” razili swoich „uczniów” prądem, to tym, co prawda, nie działa się krzywda (wstrząsu nie było), niemniej bardzo sugestywnie udawali, że faktycznie otrzymują wstrząsy elektryczne. Wielokrotnie wydawali również okrzyki bólu, prosili badanych o to, by przestali, i histerycznie krzyczeli. W pokoju z osobą badaną był również obecny eksperymentator, który w momencie, gdy „nauczyciele” mieli wątpliwości co do dalszego aplikowania wstrząsów, nakazywał im – na początku w sposób łagodny, a później już bardzo zdecydowany – kontynuować. Jeśli ktoś wyrażał wątpliwości co do bezpieczeństwa procedury, to słyszał w odpowiedzi, że no dobra, może i te wstrząsy bołą, ale „nie powodują uszkodzenia tkanek”. Co brzmi jak wymówka, którą chciałabym znać kilkanaście lat temu, gdy mama kazała mi przestać tłuc się z bratem.

Prawdą jest, że w toku badania „nauczyciele” często mieli wątpliwości, pytali, czy na pewno mają kontynuować, i przejawiali silne objawy stresu, niemniej obczajcie to: w pierwszej wersji badania 65% osób (z 40 badanych) dotarło do końca skali i zaserwowało swojemu biednemu „uczniowi” najsilniejszy szok, nawet gdy błagał on o litość i krzyczał z (udawanego) bólu⁹⁵. Raczej trudno sobie wyobrazić, by równie duża część badanych zaznaczyła w ankiecie, że „tak, raziłbym drugiego człowieka prądem, bo mi kazali”. Niekoniecznie dlatego, że celowo chcieliby skłamać, również dlatego, że jak pisała w jednym ze swoich wierszy Wisława Szymborska: „tyle wiemy o sobie, na ile nas sprawdzono”. Zauważcie, że napisała właśnie: „sprawdzono”, a nie „spytano”. I nawet nie wiadomo, czy przeprowadziła w tej kwestii jakąś ankietę.

Oczywiście eksperyment Milgrama, jak i wszystkie inne eksperymenty terenowe, nie jest pozbawiony wad. Niemniej w przypadku niektórych pytań badawczych jest to zwyczajnie bardziej wiarygodna metoda niż badania ankietowe. Osobiście uważam, że z tym konkretnym eksperymentem wiążą się dwie wiadomości. Dobra jest taka, że pozostaje

nam się cieszyć, że Milgram nigdy nie spotkał Zimbardo. Zła jest taka, że niestety nie udało mi się ustalić, jak wielu badanych w odpowiedzi na tłumaczenia, że te porażenia prądem to oni zdawali tylko dlatego, że im kazano, usłyszało tekst równie klasyczny co skuteczny, czyli: A JAKBY EKSPERYMENTATOR MAREK KAZAŁ CI Z OKNA SKOCZYĆ, TO TEŻ BYŚ SKOCZYŁ?!⁹⁶

A to nie jedyny problem, jaki badacze mają z eksperymentami terenowymi. Coraz częściej dyskutowana jest również ich etyczność (zobacz na przykład Teele 2014, West i Gunn 1978, Riach i Rich 2004), w przypadku eksperymentów ze sfalszowanymi życiorysami stawka wydaje się niewielka – zmarnowany czas pracodawców i decepcja. W takich wypadkach łatwiej nam sobie zrationalizować, że cel uświęca środki⁹⁷. Ale co z eksperymentami Milgrama, Zimbardo i innych? Takimi, które mogą spowodować u badanych realne szkody psychologiczne? Czy w takim wypadku również warto? Słuchajcie, mam odpowiedź, która niekoniecznie wam się spodoba, ale za kilka akapitów z pewnością przyznacie mi rację. Brzmi ona: zawsze mogło być gorzej. Co udowodniła swego czasu nauczycielka w szkole podstawowej, Jane Elliott.

Dzień dobry, dzieci, na dzisiejszej lekcji się podyskryminujemy

Przenieśmy się na chwilę do Stanów Zjednoczonych z lat 60., a więc czasów słusznie minionych, gdy panowała głęboka segregacja rasowa. Dokładniej rzecz ujmując, przenieśmy się do 5 kwietnia 1968 roku, czyli do dnia po zabójstwie Martina Luthera Kinga juniora. Wtedy też pewna nauczycielka szkoły podstawowej, Jane Elliott, przeprowadziła ze swoimi uczniami (a może raczej – na swoich uczniach?) następujący eksperyment⁹⁸: pierwszego dnia oznajmiła dzieciom, że niebieskookie dzieci są lepsze od brązowookich, co wynika z różnego poziomu melaniny w ich organizmach, która to wpływa na wyższą inteligencję i zdolność uczenia się. By łatwo było odróżnić od siebie obie grupy, kazała niebieskookim dzieciom zawiązać brązowe apaszki na szyjach pozostałych dzieci. Następnie niebieskookie dzieci otrzymały dodatkowe przywileje, takie jak korzystanie z placu zabaw, dodatkowe pięć minut przerwy czy też – totalnie najważniejsze – możliwość brania dokładek podczas obiadu. Ponadto Jane Elliott posadziła „lepszą” grupę w pierwszych ławkach, a brązowookie dzieci wysłała na koniec sali⁹⁹. Niebieskookim kazała

ignorować pozostałe dzieci, a brązowookim zabroniła korzystania z tego samego kraniku z wodą pitną, z którego korzystała niebieskooka grupa. Na skutki nie trzeba było długo czekać, to znaczy na początku dzieci faktycznie się opierały i dość nieufnie podchodziły do słów nauczycielki, szybko jednak zaczęły się dostosowywać. Niebieskookie dzieci stały się apodyktyczne, niegrzeczne w stosunku do tych brązowookich, czasem uciekały się również do przemocy. Brązowooka grupa również się przystosowała, choć w zupełnie inny sposób – mieli problem z wykonaniem najprostszych poleceń i rozwiązywaniem najbanalniejszych zadań. Zastanawiacie się, jak wiele czasu musiało upłynąć, by zaszła taka zmiana? Tydzień. Jeden tydzień.

A na tym nie koniec. W kolejny poniedziałek Jane Elliott zjawiała się w klasie i powiedziała, że zaszła pomyłka. Ale nie taka, jak się spodziewaliście. Otóż oświadczyła dzieciom, że jednak nie miała racji, jednak jest odwrotnie i to brązowookie dzieci są lepsze niż niebieskookie, i od tamtego momentu to one zyskały wszystkie przywileje, a także mogły zawiesić tak znienawidzone przez siebie apaszki na szyjach swoich koleżanek i kolegów. Pamiętamy jednak, że ledwie tydzień wcześniej same padły ofiarą dyskryminacji i przemocy ze strony pozostałych dzieci, można by więc pomyśleć, że w przypadku odwrócenia sytuacji same będą traktować drugą grupę lepiej i z większą sympatią. Niestety nie. Sytuacja się powtórzyła.

W licznych wywiadach Jane Elliott tłumaczyła, że *Blue eyes, brown eyes experiment* miał pokazać jej uczniom, wśród których nie było ani jednego dziecka o innym kolorze skóry, czym są rasizm, segregacja i dyskryminacja. I że nie byłaby im w stanie wyjaśnić, jak to jest być dyskryminowanym, więc dzieci musiały to poczuć. Prawdą jest, że ów eksperyment nauczył nas sporo o naszej zdolności do dyskryminacji, co trudno byłoby zbadać inaczej. Myślę jednak, że nikt nie miałby problemu z tym, by w trzydzieści sekund wymienić minimum pięć elementów, które nie były w porządku (eufemizm) pod kątem etycznym (podpowiem wam: eksperymenty na dzieciach, brak zgody rodziców, kłamstwo, zgoda na przemoc, narażenie dzieci na duży stres, a może nawet i traumę). Ostatniego dnia eksperymentu Jane Elliott posadziła wszystkie dzieci obok siebie, wyjaśniła im, że ich okłamała, że tak naprawdę wszyscy ludzie są równi, a następnie dzieci chórem obiecały, że nie będą już nigdy nikogo dyskryminować, po czym się przytulały. Jane Elliott do dziś twierdzi, że w tym wypadku cel

uświęca środki. Dzieci zaś, nawet piętnaście lat po tym doświadczeniu, doskonale pamiętały upokorzenie i traumę, jakie przeżyły w trakcie tygodnia, kiedy były dyskryminowane. Skąd to wiemy?

Stąd, że Jane Elliott sfilmowała jedną z wersji swojego eksperymentu, a także – wywiady z tymi samymi uczniami 15 lat później¹⁰⁰. Jest coś szalenie niepokojącego w tym, że nawet tyle lat po tym doświadczeniu dzieci pamiętały uczucia, które im wtedy towarzyszyły. Jest również coś zaskakującego w tym, że uważały... że totalnie było warto, a w tym wypadku cel zdecydowanie uświęcił środki.

Pozostawiam wam decyzję, co myśleć o samym eksperymencie, jak i o jego wynikach. To wymaga ogromnej dozy krytycznego myślenia, dlatego na swoich zajęciach poświęcam na tę dyskusję kolejne cztery godziny. To znaczy, o ile przed budynkiem nie leży akurat jakiś bezgłowy ptak.

Czy jak wsadzimy głowę do lodówki, a nogi do piekarnika, to
będzie nam w sam raz?

Nie.

No dobra, to może inne pytanie, bo poza wszystkim to to już było.

Ile trzeba zjeść czekolady, żeby dostać Nobla?

Patrzcie, jak daleko zaszliśmy już w naszej wędrówce po metodologicznych lasach, statystycznych wrzosowiskach i bagnach dowodów anegdotycznych! Wiemy już, jak dobrać próbę, technikę badawczą czy metodę wizualizacji. Teraz czas na najtrudniejszą część każdego badania – wyciągnięcie wniosków. Najtrudniejszą, bo tutaj największą rolę odgrywają ludzie i ich umiejętności poznawcze, a my co prawda mamy wiele wspaniałych umiejętności (taniec, śpiew, umiejętność grabienia liści), niemniej nie jesteśmy też pozbawieni wad, które w kontekście badań naukowych najczęściej przyjmują formę błędów poznawczych. O czym opowie nam Franz Messerli, człowiek rozmiłowany w czekoladzie (to możemy zrozumieć) i w Nagrodach Nobla (to możemy zrozumieć nawet bardziej, bo za Nobla można kupić dużo czekolady).

Franz Messerli był lekarzem, a przy tym lubił węglowodany, postanowił więc swego czasu obmyślić plan finansowania tej miłości do produktów o wysokiej zawartości cukru. Wymyślił to tak: może kiedyś człowiek je czekoladę, to ma to pozytywny wpływ na wszelkie procesy kognitywne, a jeśli się takiemu człowiekowi szybciej i lepiej myśli, to tym samym ma większe szanse na otrzymanie Nagrody Nobla? Plan, który miał udowodnić, że istnieje związek między spożyciem czekolady a wysoką nagrodą pieniężną, właściwie nie miał wad. No dajcie spokój, wszyscy kiedyś mieliśmy głupie pomysły po pijaku. Z tym że nie wszyscy publikowaliśmy je potem w międzynarodowych czasopismach. A Messerli tak:

Chocolate Consumption, Cognitive Function, and Nobel Laureates

Franz H. Messerli, M.D.

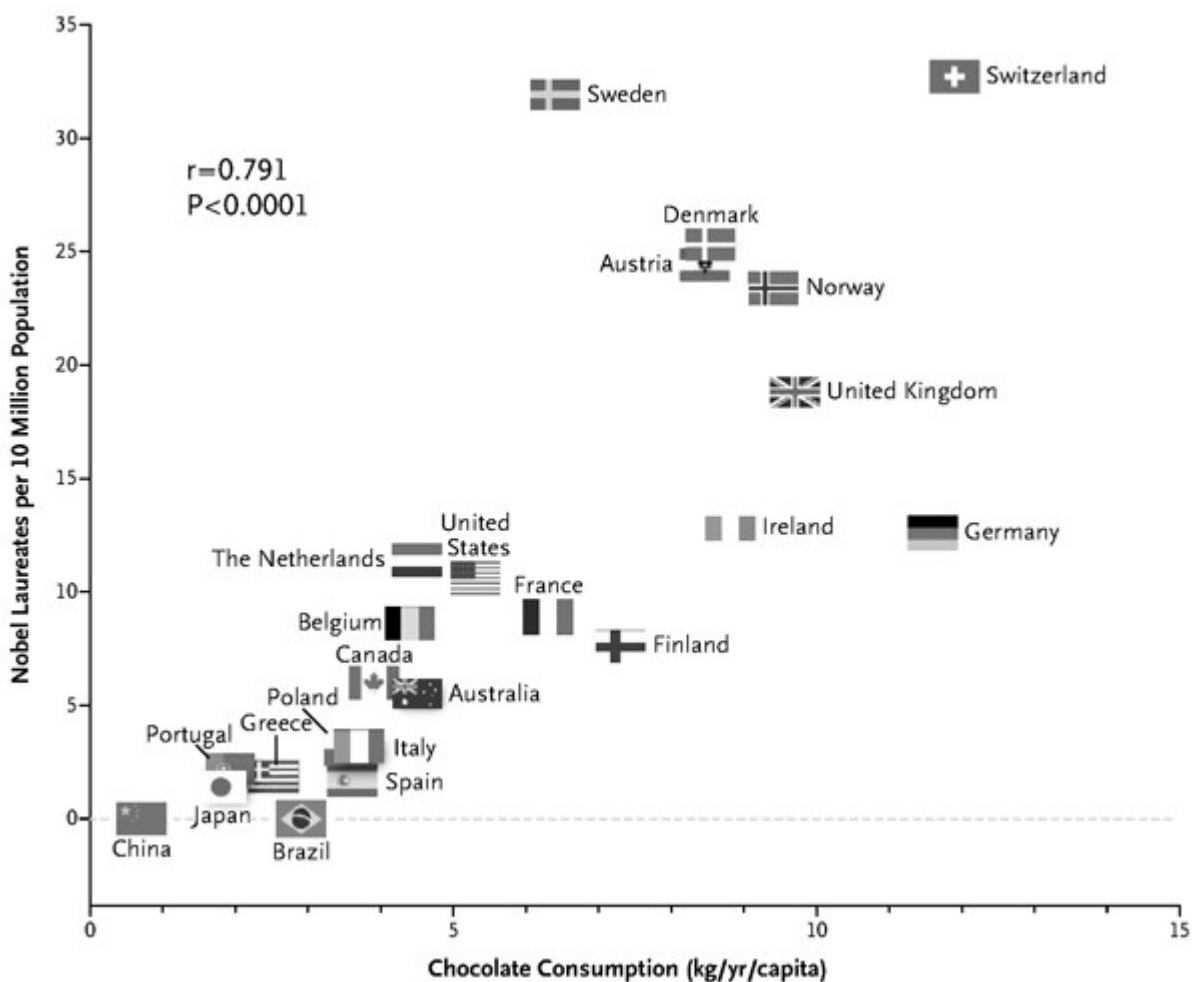
Tak właściwie to klub miłośników węglowodanów, zdeterminowanych, by znaleźć naukowe dowody na konieczność jedzenia miłych rzeczy, jest całkiem spory. Na przykład Eha Nurk (2009) wraz z grupą norweskich badaczy mierzyła za pomocą różnych testów funkcje poznawcze i sprawność intelektualną swoich badanych, a następnie pytała ich, czy często umilają sobie życie takimi radościami jak: czekolada, wino i herbata.

Tak, wiem, też nie rozumiem, co tam robi herbata. Wyniki okazały się rozkoszne: ludzie, którzy deklarowali spożywanie choć jednego z tych produktów, osiągnęli lepsze wyniki na testach niż ci, których życie pozbawione było przyjemności. A w ogóle najwyższe wyniki uzyskali ci, którzy otulali się czule pluszem czekolady, zapijali winem i jeszcze czasem herbatą, łohohoho, ale radość!

A wiecie, kto nie miał tyle szczęścia co skandynawscy badani? Szczury, które w innym badaniu były karmione ekstraktem z kakao. Czaicie, ekstraktem z kakao, to dopiero rozczarowanie! Weźcie to sobie w ogóle wyobraźcie: idziecie do restauracji, a tam podają wam ekstrakt z kotleta. Nie wiadomo, dlaczego Jean-Francois Bisson (2008) poskąpił swoim puchatym badanym prawdziwej czekolady, może bał się, że one mają takie małe łapki, że wieki całe miną, zanim odpakują choćby jedną tabliczkę. Niemniej wyniki były obiecujące – szczury karmione kakaowym ekstraktem przechodziły przez labirynt znacznie szybciej niż ich koledzy na diecie.

Zainspirowany swoimi poprzednikami główny bohater naszej historii, Franz Messerli, postanowił wznieść swoje badania na trochę wyższy poziom, dokładniej rzecz ujmując: na poziom międzynarodowy. I wymyślił sobie tak: weźmie kilka krajów, sprawdzi, jak się w nich miewa spożycie czekolady, a także ilu laureatów Nagrody Nobla można w owym kraju znaleźć, a jeśli się okaże, że w krajach, gdzie ludzie kąpią się w czekoladzie, jest więcej noblistów, no to ja was bardzo przepraszam, ale mnie nie trzeba tego dwa razy powtarzać, nie pozostanie nam nic innego jak lecieć do najbliższego sklepu.

Jak postanowił, tak zrobił. Zebrał odpowiednie dane, a następnie naszkicował je na wykresie rozrzutu. Na osi X wyrysował średnie spożycie czekolady w danym kraju, w kilogramach na osobę. Na osi Y zaś – liczbę noblistów w danym kraju (od początku przyznawania nagród do 2010 roku). Następnie zaś naniósł na ten graf poszczególne kraje ¹⁰¹.



No i patrzcie: Chiny – bardzo niskie spożycie czekolady i żadnych nagród. Tacy byli zajęci tym, żeby zostać potęgą gospodarczą, że zapomnieli starać się o Nobla. Sporo krajów jest gdzieś tam w połowie stawki, zarówno jeśli chodzi o spożycie czekolady, jak i liczbę noblistów. Ale patrzcie na naszą gwiazdę – Szwajcaria. Największe spożycie czekolady i największa liczba noblistów. KABOOM. Dziękuję, można się rozejść do półek sklepowych.

To, co widzimy na wykresie, nazywa się **KORELACJĄ POZYTYWNA**. Występuje ona wtedy, kiedy wraz ze wzrostem jednej zmiennej rośnie wartość drugiej zmiennej (i na odwrót). Na przykład: im częściej rozmawiam przez telefon, tym wyższe rachunki płacę. Im więcej się uczę, tym lepsze mam oceny. A w tym wypadku: im większe spożycie czekolady w danym kraju, tym większa liczba noblistów. Przeciwnościem korelacji pozytywnej jest **KORELACJA NEGATYWNA**, czyli taka, w której wartości obu cech postępują w kierunkach przeciwnych – gdy wartość jednej zmiennej rośnie, to wartość drugiej maleje. Przykładowo: im jestem

starszy, tym mam gorsze zdrowie. Im więcej czasu spędzam w pracy, tym mniej czasu spędzam z rodziną.

Oprócz kierunku korelacji statystyka pozwala nam też zmierzyć jej siłę. Żeby to zrobić, musimy policzyć współczynnik korelacji (w tym wypadku współczynnik korelacji Pearsona¹⁰²), który przyjmuje wartości od -1 do 1 i oznaczamy go literką R. Ja lubię myśleć, że ta literka R oznacza „rozkosz”, moi studenci zaś uważają, że to bardziej takie R jak „rany boskie”.

Wartość -1 współczynnika korelacji oznacza idealną korelację negatywną, wartość +1 idealną korelację pozytywną. 0 oznacza absolutny brak korelacji.

Ej, a byliście kiedyś na walce bokserskiej? Jeśli chcecie takową zobaczyć, to wystarczy, że spytacie dwoje badaczy o to, jaka wartość współczynnika R oznacza silną korelację. Otóż ilu badaczy, tyle opinii, a dla wielu z nich jest to temat na tyle istotny, że zasługuje na walkę wręcz. Choć oczywiście mówimy o bitwie dwóch naukowców, więc to trochę tak, jakby obserwować dwójkę niemowlaków obrzucających się watą. Tyle że niemowlaki bywają w tym wszystkim skuteczne.

Na przykład kolega Guilford (1954) uważał, że współczynnik korelacji powyżej 0,5 oznacza silną korelację pozytywną. Jego kolega Evans (1996) był znacznie surowszy i za silną korelację uznał współczynniki powyżej 0,6. To, jaką klasyfikację wybierze, zależy trochę od was, trochę od waszej dyscypliny (w sensie dyscypliny naukowej, a nie takiej życiowej, bo tej drugiej to ja na przykład w ogóle nie posiadam, wie to każdy, kto widział mnie kiedyś przy bufecie śniadaniowym w hotelu). Nauki ścisłe i medyczne są w tej kwestii znacznie bardziej wymagające niż nauki społeczne.

Niemniej nasz główny bohater, Franz Messerli, w ogóle nie musiał się klasyfikacjami przejmować, albowiem jego współczynnik określający siłę związku między jedzeniem czekolady a Noblami był według wszelkich interpretacji sympatycznie wysoki i wynosił $R = 0,791$. Tego nie trzeba było Franzowi dwa razy powtarzać – natychmiast wysmarował szalenie optymistyczne wnioski, które w wolnym tłumaczeniu brzmiały tak: „Jedzenie czekolady wzmacnia funkcje poznawcze, które są warunkiem koniecznym do otrzymania Nagrody Nobla, i ściśle współwystępuje z liczbą laureatów tej nagrody w każdym z krajów”¹⁰³.

To może ja powtórzę: „jedzenie czekolady wzmacnia funkcje poznawcze, które są warunkiem koniecznym do otrzymania Nagrody Nobla”. Wiecie, gdzie jeszcze znajdziecie to zdanie? Na karcie oficjalnego skierowania Messerlego do piekła, gdzie będzie się smażył do końca życia wraz z pozostałymi ludźmi, którym zdarzyło się pomylić korelację z przyczynowością. A to jest bardzo brutalny krąg piekła, bo on sąsiaduje z kręgiem przeznaczonym dla ludzi, którzy za życia odgrzewali rybę w służbowej mikrofalówce, i wiecie co? Po śmierci też się nie mogą powstrzymać.



Słuchajcie, nie ma tylu wykrzykników na tym świecie, by wyrazić moją stanowczość – korelacja nigdy nie jest równoznaczna ze związkiem przyczynowo-skutkowym!!! Nigdy. Nie jest również prawdą – choć jest to szalenie popularny mit – że wystarczająco silna korelacja pozwala uznać, że jeden czynnik wpływa na drugi. Znana jest historia Romana z Bytomia, który w 1992 podzielił się z wszystkimi taką właśnie ciekawostką: „Hej – powiedział – czy wiecie, że wystarczająco silna korelacja oznacza związek przyczynowo-skutkowy?”, a chwilę później za karę dostał od losu najmniejszy kawałek sernika.

Co może oznaczać korelacja?

Korelacja nie musi oznaczać związku przyczynowo-skutkowego, choć czasem rzeczywiście taka przyczynowość istnieje – ja tam nie potrzebuję

absolutnie żadnych badań, żeby wiedzieć, że jedzenie czekolady poprawia jakość życia i z każdą dodatkową kostką człowiekowi robi się w środku znacznie bardziej pluszowo. Niemniej w pozostałych kwestiach, gdy odkryjemy korelację między czynnikiem A i czynnikiem B, to możemy mieć do czynienia z jednym z pięciu podstawowych przypadków:

1. ZMIENNA A RZECZYWIŚCIE WPŁYWA NA ZMIENNĄ B, tak jak w przypadku długości rozmów telefonicznych i późniejszych rachunków albo liczby kostek czekolady i poczucia pluszowatości.

2. ZMIENNA B WPŁYWA NA ZMIENNĄ A, czyli zależność odwrotna. Takie ustalenie kierunku korelacji nie zawsze jest intuicyjne. Czasem tak – na przykład w przypadku silnej korelacji między padającym deszczem a używaniem parasoli. Nikt nie ma wątpliwości, że to obecność deszczu powoduje, że ludzie zaczynają szybko otwierać parasole, a nie na odwrót. Często jednak ustalenie, co wpływa na co, jest znacznie bardziej skomplikowane. W 2007 roku dr Shirli Werner wraz ze współpracownikami przeprowadziła badanie, w którym wykazała, że schizofrenia występuje częściej wśród ludzi o niższym statusie społeczno-ekonomicznym (Werner 2007) – ta narracja latami była dominująca w psychiatrii i epidemiologii. Niemniej doktor B. Burt Gerstman (2013) w swoim podręczniku do epidemiologii zasugerował, że możemy mieć tu do czynienia z ODWRÓCONĄ PRZYCZYNOWOŚCIĄ (ang. *reversed causation*), czyli że to nie niski status społeczno-ekonomiczny powoduje schizofrenię, ale sama choroba może być odpowiedzialna za problemy finansowe i pogorszenie stopy życiowej chorych.

3. ZMIENNA A WPŁYWA NA ZMIENNĄ B I JEDNOCZEŚNIE ZMIENNA B WPŁYWA NA ZMIENNĄ A. Taka karuzela przyczynowości mnie osobiście jara znacznie bardziej niż klasyczne karuzele, w których człowiek wsiada do plastikowego łabędzia, a potem przez pół godziny kręci się w kółko jak idiota. Przykład? Sprzedaż krasnali ogrodowych i budżet reklamowy. Im większe nakłady na reklamę, tym większa sprzedaż krasnali, a im lepiej kręci się sprzedaż krasnali, tym więcej pieniędzy mamy na kolejną reklamę.

4. ISTNIEJE TRZECIA ZMIENNA C, KTÓRA KORELUJE JEDNOCZEŚNIE ZE ZMIENNYMI A I B. Czy wiecie, że istnieje bardzo silna i wielokrotnie udowodniona korelacja między liczbą sprzedanych lodów a... liczbą utonięć? Ale czy to oznacza, że od dziś musicie jeść różka truskawkowego w kapoku, a do pobliskiej lodziarni na wszelki wypadek

chodzić z własnym pontonem? No nie. Istnieje inne wyjaśnienie tak silnego związku i jest to oddziaływanie trzeciej zmiennej – wysokich temperatur. Gdy jest lato, to ludzie częściej kupują lody, ale również – częściej wchodzi do wody (niezależnie od tego, czy jedli lody, czy też nie) i niestety czasem toną.

5. KORELACJA JEST PRZYPADKOWA. Serioszka, po prostu zdarza się przez zupełny przypadek, zupełnie jak moja ostatnia wizyta na siłowni. Z szalonych, aczkolwiek totalnie przypadkowych korelacji słynie między innymi strona internetowa *Spurious correlations*. Tam możemy zobaczyć, że istnieje na przykład bardzo silna korelacja między konsumpcją sera a ryzykiem śmierci poprzez zasypianie w pościeli, co w sumie brzmi bardzo groźnie, ale mam przeczucie graniczące z pewnością, że to nie jest tak, że jak człowiek sobie strzeli przed snem 100 gramów goudy, to następnie padnie ofiarą śmiertelnego potrzasku w prześcieradle.

Więc skoro już wiemy, że nie musimy przechodzić na drugą stronę chodnika, gdy napotkamy na ulicy kostkę żółtego sera, to zastanówmy się, co z tą czekoladą.

Jeśli nie korelacja, to co?

Badanie Messerlego wywołało całkiem uzasadnione poruszenie, zarówno wśród naukowców, jak i normalnych, szczęśliwych ludzi. Zanim jednak przyczaimy się na najdoskonalszą promocję, a następnie zaczniemy ćwiczyć mowę, w której podziękujemy mamie, tacie i wszystkim członkom Akademii, warto zastanowić się nad tym, co jeszcze może mieć znaczenie dla liczby noblistów w danym kraju? Jednym z czynników, który przychodzi do głowy chyba wszystkim i przyszedł też do głowy psychologowi Pierre'owi Maurage'owi (2013) wraz z naukową drużyną – był stan finansowy danego państwa. Może jest tak, że jeśli kraj jest bogatszy, to przeznaczają więcej wiader złotych monet na naukę, infrastrukturę naukową, stypendia, a życie naukowców w dobrobycie jest całkiem miłą pożywką dla potencjalnych Nagród Nobla. Maurage poszedł nawet dalej – pomyślał, że PKB, podstawowy wskaźnik ekonomiczny używany do oceny stanu danej gospodarki, może mieć znaczenie zarówno dla liczby noblistów w danym kraju, jak i dla spożycia czekolady.

Żeby to sprawdzić, skorzystałam z najwspanialszego odkrycia ludzkości zaraz po penicylinie, kole i maśle orzechowym – regresji statystycznej¹⁰⁴. Regresja jest to pewien model matematyczny, który pozwala eksplorować związki pomiędzy więcej niż dwiema zmiennymi. To oznacza, że za pomocą regresji mogłam sprawdzić nie tylko czy istnieje związek między jedzeniem czekolady a Noblem, lecz także dodać do tego związku miłosnego pozostałe osoby dramatu, czyli takie zmienne jak PKB czy temperatura w danym kraju (która pewnie nie ma zbyt dużego znaczenia dla noblistów, ale dla spożycia czekolady już tak, bo wyobraźcie sobie, że idzie taki człowiek ulicami jakiegoś ekstremalnie ciepłego kraju, kupuje czekoladę, dochodzi do chałupy, a tam kakao).

Jak sobie człowiek trzaśnie regresję, to oprócz tego, że jakby dzień od razu jest lepszy, to jeszcze dostaje nagrodę w postaci współczynników regresji przyporządkowanych każdej zmiennej (widzimy je w tabeli poniżej).

OD CZEGO ZALEŻY LICZBA NOBLISTÓW? – TABELA REGRESJI

Czekolada/osobę (w kilogramach)	0,014
PKB/osobę	0,034**
Średnia temperatura (w stopniach Celsjusza)	-0,04**

Jeśli chodzi o kierunek (dodatni lub ujemny) współczynnika, to interpretujemy go podobnie jak współczynnik korelacji. Trochę inaczej jest z interpretacją jego wielkości. Współczynnik regresji określa nam, o ile wzrośnie wartość zmiennej zależnej (w tym wypadku liczba noblistów), gdy wartość zmiennej niezależnej (w tym wypadku spożycie czekolady) wzrośnie o 1, a wartość pozostałych zmiennych pozostanie taka sama. W przypadku naszego modelu interpretujemy ten współczynnik w taki sposób: gdyby absolutnie każdy człowiek w danym kraju zjadł dodatkowy kilogram czekolady w ciągu roku, to przewidywalnie w zamian otrzymalibyśmy 0,014 noblisty. No słuchajcie, nie trzeba być geniuszem matematyki, żeby wiedzieć, że 0,014 człowieka to niezbyt dużo, rzekłabym: czubek głowy. I ja bym nie chciała tutaj dyskryminować różnych części ciała, uważam, że rola czubka głowy jest szalenie istotna,

niemniej nie da się ukryć, że w kwestii liczby noblistów to znaczenie czekolady jest w gruncie rzeczy bardzo niewielkie.

Zarówno w moim modelu matematycznym, jak i w tym, który zastosował Pierre Maurage wraz z kolegami, wnioski były takie same – korelacja między spożyciem czekolady a liczbą noblistów jest pozorna. Znacznie większe znaczenie dla tego, jaką liczbą laureatów państwo będzie się mogło cieszyć, ma jego stan finansowy, a nie to, czy zjemy więcej czekolady. Co oczywiście nie oznacza, że nie powinniśmy próbować.

Jak jeszcze pomylił się Messerli?

Nie mam żalu do Messerlego, że zbytnia miłość do czekolady przysłoniła mu rozsądek i podczas pisania wniosków trochę odjechał mu peron. Niemniej nie jest to jedyny błąd metodologiczny, który popełnił w swoim artykule. Po pierwsze, wiecie, w jaki sposób Messerli zmierzył spożycie czekolady w danym kraju? Czy stał koło półki ze słodyczami w Tesco i skrupulatnie liczył, kto kupuje czekoladki? No nie. Czy poszedł do Światowego Ministerstwa Węglowodanów i poprosił ich o oficjalne dane? No też nie, bo takie ministerstwo nie istnieje, co osobiście uważam za bardzo duże niedopatrzenie, bo już się szykowałam na ciepłą posadę. Otóż Messerli wziął dane z bardzo konkretnego źródła – od producenta konkretnej marki czekolady.

Uczcijmy to minutą metodologicznej ciszy.

Innymi słowy, nasz główny bohater wcale nie mierzył spożycia czekolady w danym kraju. Mierzył spożycie tylko i wyłącznie gorzkiej czekolady marki Lindt. Marki, notabene, szwajcarskiej, co może częściowo tłumaczyć, dlaczego jej spożycie jest najwyższe właśnie w tym kraju.

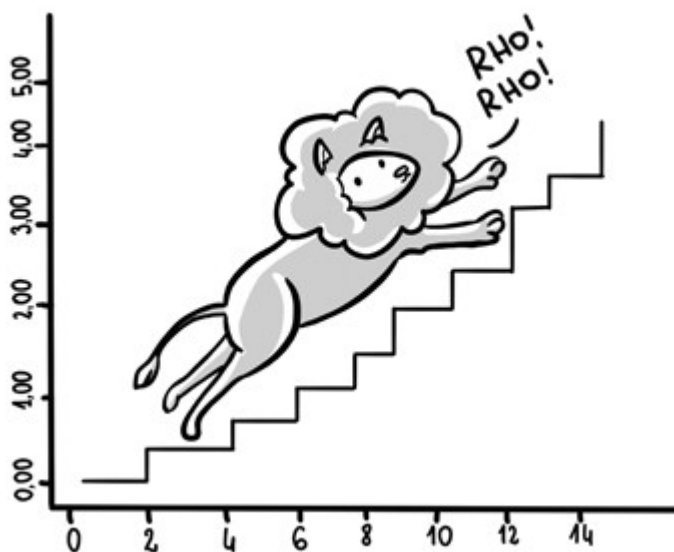
To jeszcze nie wszystko, może być tak, że Messerlemu ścielą czułe gniazdko w metodologicznym kręgu piekła jeszcze z jednego powodu – badacz ten mierzył związek między średnim spożyciem czekolady w danym kraju a liczbą noblistów, a więc ustanowił związek na poziomie krajowym. Niemniej na tej podstawie wysnuł wnioski o zależnościach jednostkowych. To znaczy uznał, że to, że w krajach, w których je się więcej czekolady, jest więcej noblistów, jest równoznaczne z tym, że ludzie jedzący czekoladę będą mądrzejsi. To bardzo popularny błąd. Właściwie to tak popularny, że ma nawet swoją nazwę. Nazywa się **BŁĘDEM EKOLOGICZNYM ROZUMOWANIA** (ang. *ecological fallacy*) i dowiemy się o nim więcej

w kolejnym rozdziale. Polega on na niepoprawnym przenoszeniu wniosków z korelacji lub obserwacji grupowych (np. na poziomie kraju, społeczności czy mieszkańców danego miasta) na jednostkę wchodzącą w skład tej grupy. To znaczy: to, że ustanowiono związek między bogactwem danego kraju a przeciętnym stanem zdrowia jego obywateli, wcale nie oznacza, że bogaci ludzie będą żyć wiecznie, a przynajmniej długo i szczęśliwie. Później wrócimy do tego przykładu.

Analogicznie nie jest prawdą, że pozytywna korelacja między spożyciem czekolady a liczbą noblistów oznacza, że jeśli teraz zaczniemy się nadziewać czekoladą, to potem pozostanie nam już tylko czekać na telefon ze Sztokholmu. Ale nie traćmy nadziei! Skoro dzięki jedzeniu czekolady nie dostaniemy Nobla, to może chociaż kąpiele w maśle zapewnią nam nieśmiertelność?

Czy język angielski powoduje zawały serca, a masło przedłuża życie?

Hans Rosling, wspomniany już wcześniej najwspanialszy statystyk świata, opowiadał w jednym ze swoich licznych i oszałamiających wystąpień anegdotkę o korelacji. Idea anegdotek o kwestiach statystycznych jest mi bardzo bliska, bo sama również lubię takowe od czasu do czasu uskuteczniać. Na przykład istnieje sobie w statystyce taki współczynnik Rho i w sumie to w ogóle nie jest istotne, co on robi, najważniejsze jest to, że ja go sobie ostatnio liczyłam i jak już policzyłam, to przy okazji wymyśliłam brawurowy dowcip, ale trzymajcie się krzeseł, bo on jest taki zabawny, że zaraz wszyscy pospadacie. Bo wiecie normalnie, kiedy angielskojęzyczne lwy chodzą po świecie, to ryczą: *Roar! Roar!*, a mój dowcip był taki, że uwaga: wiecie, jak ryczy lew, który lubi statystykę? Otóż odpowiadam: *Rho! Rho!*

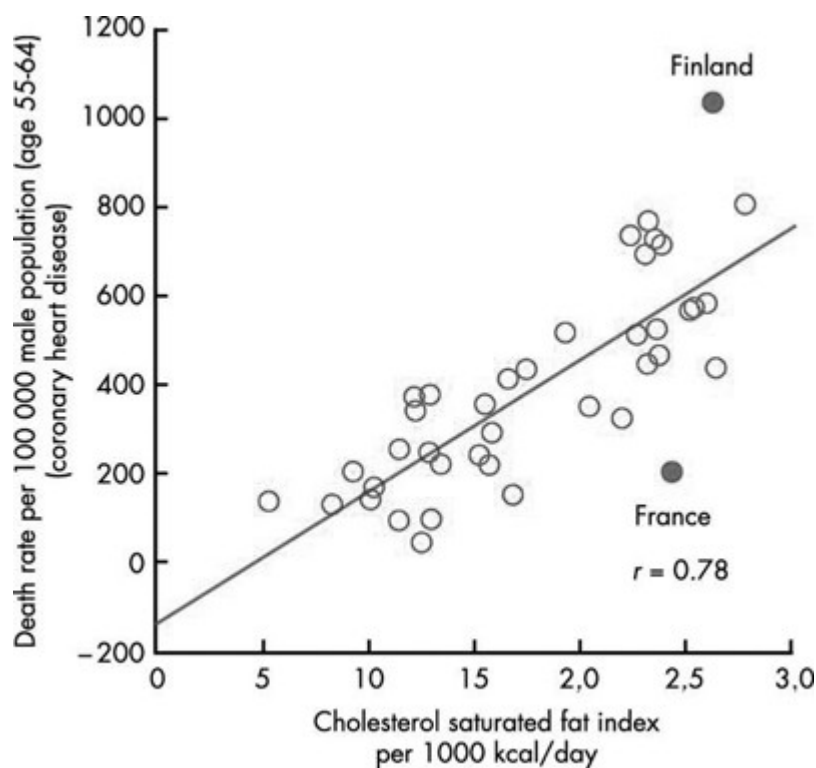


O mój Boże, w mojej głowie to było takie zabawne, że jak to wymyśliłam, to trzy godziny płakałam ze śmiechu, a potem poszłam do kolegi z ekonometrii i spytałam go, czy wie, jak ryczy lew, który lubi statystykę, a on nie wiedział, więc mu powiedziałam, że *Rho! Rho!*, a on powiedział, że to przezabawne, i krzyknął do kolegi spod okna, czy wie, jak ryczy lew, który lubi statystykę, a tamten też nie wiedział, więc on mu odkrzyknął, że ten lew to robi: *Rho! Rho!*, i ten kolega też się zaczął

strasznie śmiać, a potem jeszcze chcieliśmy opowiedzieć ten żart reszcie naszych przyjaciół, ale nie mogliśmy, bo przypomnieliśmy sobie, że nie mamy żadnych przyjaciół. I w sumie powoli zaczynam rozumieć dlaczego.

Niemniej Hans Rosling na pewno śmiałby się z mojej anegdotki, jak i ja się śmiałam z jego. Opowiadał tak: pewien miły człowiek zastanawiał się, jaki jest związek między piciem alkoholu oraz tłustym jedzeniem a zawałami serca. Wymyślił sobie, że w Japonii piją mało alkoholu, jedzą sporo tłustych ryb i mają mało zawałów serca. Następnie sprawdził, co słychać u Francuzów, i odkrył, że ci znów piją dużo alkoholu, jedzą dużo tłustych serów i też mają mało zawałów serca. Potem popatrzył na Wielką Brytanię – piją stosunkowo niewiele, ale jedzą dość tłusto, mają dużo zawałów serca. Zanotował w dzienniczku badawczym: „wygląda na to, że przyczyną zawałów serca jest język angielski”.

Niektórzy (Francuzi) lubią tę anegdotkę aż za bardzo, na tyle, że nadali jej swoją nazwę. W 1981 roku trzech Francuzi: Ducimetière, Cambien i Richard, stworzyli nazwę *French paradox* na opisanie tego dziwnego zjawiska, w którym pomimo tłustego jedzenia i ogromnych ilości alkoholu ich mili rodacy w beretach cieszą się niezwykle dobrym zdrowiem. Nazwa została stworzona po tym, jak porównali dane dotyczące zawałów serca i spożycia tłuszczu w poszczególnych krajach¹⁰⁵. Dane te możemy zobaczyć na wykresie na następnej stronie:



Na tym wykresie rozrzutu widzimy dwie zmienne: na osi X – spożycie cholesterolu (na 1000 kcal/dziennie), na osi Y – śmiertelność wskutek chorób serca (na 100 000 obywateli płci męskiej w wieku 55–64). Każda kropka reprezentuje jedną obserwację, czyli w tym wypadku jeden kraj. Te śmieszne uszczegółowienia typu „śmiertelność na 100 000 obywateli” czy „cholesterol na 1000 kcal/dziennie” są używane w celu standaryzacji, tak byśmy mogli porównywać między sobą kraje o populacjach różnej wielkości i te przyjmujące różną dzienną dawkę kalorii¹⁰⁶.

Jak widać na wykresie, istnieje korelacja pozytywna między spożyciem tłuszczów a śmiercią spowodowaną chorobami serca, co oznacza, że im większe spożycie tłuszczów w danym kraju, tym większa śmiertelność. Niemniej istnieją dwa kraje, które nie pasują do tego trendu. Pierwszym jest Finlandia, która ma znacznie większą śmiertelność, niż wskazywałby na to poziom spożywanego tłuszczu. Próbowano to wyjaśnić wyższym w tym kraju spożyciem tłuszczu pochodzącego z mleka, śmietany, sera i masła i mniejszym spożyciem warzyw. Niemniej jeśli ktoś czyta tę książkę przy zupie zrobionej z masła i śmietany, przyprószonej kruszonką z goudy, to śpieszę go uspokoić, że ta hipoteza nie została nigdy zbadana i potwierdzona. Głównie dlatego, że największą uwagę badaczy zwrócili ci szczęściarze, którzy taplają się w maśle, a mimo to mają zdrowe serca – Francuzi.

Czy taplanie się w maśle naprawdę wydłuża życie?

Badacze mieli różne pomysły na to, jak rozwiązać tę zagadkę. Wielu z nich twierdziło, że za taki pomyślny dla Francuzów stan rzeczy odpowiada resweratol, substancja należąca do grupy antyoksydantów, która występuje w znacznych ilościach w czerwonym winie, którego Francuzi piją najwięcej. Resweratol rzeczywiście jest substancją, która chroni rośliny przed atakiem pasożytów, bakterii i nadmierną ekspozycją na promieniowanie UV. Liczne badania na zwierzętach i ludziach potwierdzają jej wspaniałe właściwości. I ja nie potrzebuję więcej dowodów, bardzo przepraszam, ale następny rozdział może być nieco chaotyczny.

Problem polega na tym, że – według badań – ochronne działanie resweratolu zaczyna się od dawki 500 mg. Żeby takową dawkę spożyć,

musielibyśmy wypić 40 litrów wina na głowę, a umówmy się, nie wszyscy mieszkamy w akademiku.

Były również podejrzenia, że chętnie jedzone przez Francuzów warzywa równoważą złe działanie tłuszczów (Artaus-Wild et al. 1993), z czego ucieszyła się głównie brukselka, wobec której od wielu lat prowadzona jest agresywna polityka nienawiści. Niektórzy wskazywali też na to, że francuscy lekarze mieli inny system klasyfikowania zgonów, również tych spowodowanych chorobami serca – niemniej te odmienności w klasyfikacjach okazały się wyjaśniać różnice pomiędzy krajami jedynie w niewielkim stopniu (Law i Wald 1999). Warto jednak pamiętać o tym, że wysoki poziom cholesterolu jest pewną bombą z opóźnionym zapłonem (podobnie jak szkodliwość papierosów czy alkoholu) – konsumpcja tłuszczu i tym samym średni poziom cholesterolu wzrosły we Francji znacznie później niż w pozostałych krajach, mniej więcej 15 lat od pierwszego, wciąż licznie cytowanego, badania. Naukowcy Law i Wald (1999) słusznie zwrócili uwagę, że opóźnienie czasowe między nadużywaniem tłuszczu a negatywnymi objawami wynosi 25 lat; zamiast więc wciąż zazdrościć Francuzom ich romansów ze śmietaną, warto byłoby to badanie powtórzyć, by sprawdzić, jak się ma stan kardiologiczny różnych narodowości obecnie.

Przede wszystkim jednak cała ta dyskusja – choć moim zdaniem jak najbardziej uzasadniona, no bo kto z nas nie chciałby się kąpać w maśle niczym najwspanialszy w świecie pieróg – była przykładem bardzo konkretnego błędu, który już zresztą znacie. Tak jest, mowa o wspomnianym wcześniej istnym eklerku wśród błędów wnioskowania, najpuchatszej z owiec tego statystycznego pastwiska, błędzie tak wspaniałym, że gdyby kiedyś wystąpił w programie *Rolnik szuka żony*, toby totalnie jakąś znalazł. Błędzie ekologicznym rozumowania.

Z dobrych wiadomości: jeśli zdarzyło wam się kiedyś ten błąd popełnić, to jesteście w doborowym towarzystwie: Prezydenta RP, najślynniejszego socjologa świata i prawie każdej sąsiadki. Ekipa, jakich mało...

Co ma wspólnego prezydent Andrzej Duda z Emizem Durkheimem?

Wszystko zaczęło się od międzynarodowego szczytu klimatycznego w Katowicach, zwanego również festiwalem ironii, który odbył się w 2018 roku. Jako człowiek, który ma serce z węgla czule wyciosane dzieciństwem

na Śląsku, z radością przyjmowałam kolejne doniesienia, jakoby poważnych przedstawicieli szczytu, przypomnę, klimatycznego, witały czułe dźwięki orkiestry górniczej, tak jest – GÓRNICZEJ. I ma to sens, bo to z pewnością dobrze komponowało się z torbami konferencyjnymi, w których – totalnie tak to sobie wyobrażam – każdy delegat mógł znaleźć symboliczną bryłkę węgla, a także zapałki i kalosza, bo to był taki polski zestaw DIY: ogrzej się zimą. Ponadto wszystkim zgromadzonym z dumą pokazywano ostatnie pomiary stężenia pyłu w powietrzu, z nadzieją, że się nie kapną i pomyślą, że to nasze PKB.

Niemniej wiadomo, jak jest – tam, gdzie dobra wiksa, tam zawsze znajdzie się jakiś oburzony Mariusz, który cały wieczór będzie sący sok z kiszonych ogórków, bo zdrowy na nerki, i pokazywał wszystkim swój klaser ze znaczkami, a potem zepsuje dobrą zabawę, puszczając *Rzekę marzeń* Beaty Kozidrak, z którą to zresztą Beatą wspólnie będzie śpiewał wszystkie refreny. No i przy okazji naszego festiwalu ironii i palonego kalosza, który odbył się w Katowicach, też oczywiście się taki znalazł. I przypomniał tweeta pana prezydenta, w którym to tweecie przeczytać można o tym, że ha, ha, ale wszyscy oszukują z tym globalnym ociepleniem, skoro za oknem jest zimno, ha, ha.



I teraz na pewno zastanawiacie się, co jest nie tak w tym stwierdzeniu, oprócz tego, że wszystko. I co festiwal palonego kalosza ma wspólnego z kąpaniem się w śmietanie. Otóż tak: rozumowanie z powyższego tweeta również jest przykładem tego samego błędu – błędu ekologicznego rozumowania.

To jest taki błąd, który polega na nieuzasadnionym przenoszeniu wniosków z korelacji lub obserwacji grupowych na jednostki wchodzące w skład tej grupy, i ja wiem, że tutaj większość z was usłyszała „bla, bla, bla”, chyba że akurat jesteście labradorem, to z tego wszystkiego zrozumieliście tylko tyle, że „jednostki” wspaniale rymuje się z „kostki”, i już jesteście przy misce. Ale tak po ludzku to chodzi o to, że jeśli istnieje

jakaś zależność wysokiego poziomu, na przykład że im bogatszy kraj, tym większa średnia długość życia jego mieszkańców, to wcale nie oznacza, że wystarczy pływać w monetach i w ten sposób staje się nieśmiertelnym. Albo że jak ktoś sobie przeczytał, że im biedniejszy kraj, tym większa średnia długość penisa (Westling 2011), to teraz warto pod koniec każdego trudnego miesiąca stać w łazience z linijką i dużą dozą nadziei.

To, że coś jest faktem na poziomie kraju, społeczności czy mieszkańców danego miasta, wcale nie oznacza, że możemy sobie tak po prostu wziąć i skopiować te wnioski na poziom indywidualny. To, że w policyjnych statystykach istnieje korelacja dotycząca koloru skóry i popełnianych przestępstw, nie oznacza, że osoby o innym kolorze skóry niż biały częściej popełniają przestępstwa.

Wyjaśnienie jest zupełnie inne (i przykre): są po prostu częściej aresztowani i skazywani na karę pozbawienia wolności¹⁰⁷.

Badacz William Robinson był jednym z pierwszych, którzy zwrócili uwagę na ten problem w metodologii, i nadał mu nazwę, której używamy do dziś. Było to w 1950 roku! W swoim oryginalnym artykule zwrócił również uwagę na to, że zależności na poziomie grupowym i indywidualnym mogą się różnić nawet kierunkiem korelacji. Dla przykładu: korelacja między urodzeniem poza granicami Stanów Zjednoczonych a analfabetyzmem wynosiła $+0,118$ na poziomie indywidualnym (Robinson użył danych z 1930 roku dla Stanów Zjednoczonych). Była więc pozytywna, co prawdopodobnie odzwierciedlało niższe standardy edukacyjne w krajach pochodzenia migrantów. Zależność na poziomie poszczególnych stanów była jednak odwrotna – korelacja między procentem migrantów w danym stanie a procentem analfabetyzmu wynosiła $-0,526$, była więc negatywna! Dlaczego? Bo imigranci najczęściej zamieszkiwali stany silnie uprzemysłowione, jako że tam było łatwiej znaleźć pracę. Równocześnie były to stany bogatsze i o niższym poziomie analfabetyzmu niż biedniejsze stany robotnicze o bardzo dużym bezrobociu, które nie przyciągały tak wielu migrantów.

Błąd ekologiczny rozumowania to błąd popełniany szczególnie często (ale nie tylko!) w epidemiologii, na przykład we wspomnianym przez nas badaniu o maśle i chorobach serca, w którym porównywano zagregowane dane na poziomie krajowym i na tym poziomie wnioskowano o czynnikach ryzyka, które oddziałują na jednostkę. Niemniej, tak jak już wspominałam,

jeśli komukolwiek się zdarzyło taki błąd wnioskowania popełnić (bez strachu, mnie też się zdarzyło), to jest w doborowym towarzystwie – nie tylko samego pana prezydenta, ale również Emila Durkheima, przez wielu uznawanego za cesarza statystyk społecznych. Co oznacza mniej więcej tyle, że jeśli ktoś kiedyś studiował jakiegokolwiek nauki społeczne, to na bank o tym panu słyszał. No chyba że nigdy nie dotarł na żaden wykład, bo w drodze na zajęcia rozkojarzyły go guziki w windzie.

No i teraz kłopotek jest taki, że za przykład błędu ekologicznego rozumowania uznaje się jedno z najśłynniejszych badań Emila Durkheima z XIX wieku – to, w którym analizował różne statystyki dotyczące samobójstw¹⁰⁸. I w tym miejscu to ja mam taką śmieszłą anegdotkę, mianowicie kiedyś spytałam szefa, o czym mogę poprowadzić swoje absolutnie pierwsze zajęcia na uczelni, a także absolutnie pierwsze zajęcia tych miłych studentów, świeżaczków takich, co to na pierwszym roku byli i w pierwszym dniu studiów, a on mi na to powiedział, że mam im opowiedzieć o tych statystykach samobójstw. No i opowiedziałam. Połowa studentów się potem z tych zajęć wypisała. Ja w sumie też chciałam, ale okazało się, że nie mogę.

No i na przykład w tym badaniu nasz przyjaciel Emil zrobił tak, że przeanalizował dane dotyczące śmierci samobójczych w pruskich prowincjach i odkrył, że istnieje silna pozytywna korelacja między samobójstwami a odsetkiem protestantów zamieszkujących dane tereny (Durkheim 1897). Czyli im więcej protestantów w danej prowincji, tym więcej samobójstw. I to jeszcze było spoko odkrycie, ale wniosek, który na tej podstawie wysnuł, że oto protestanci częściej odbierają sobie życie niż katolicy, już nie był spoko i dawno poinformowałam o tym statystyczną policję.

Bo słuchajcie, ta korelacja grupowa nie oznaczała wcale, że protestanci częściej odbierali sobie życie! Rzeczywiście, większość samobójstw była popełniana w prowincjach zasiedlonych przez protestantów, ale samobójstwa popełniała tam głównie... nieprotestancka mniejszość (najpewniej z przyczyn nieprzystosowania i wykluczenia).

KABOOM!

Ale wam sprzedałam klawą anegdotkę do opowiedzenia przy wigilijnym stole, nie?

Więc pamiętajcie, wnioski to nie postanowienia noworoczne, że je tak można cały czas kopiować i wklejać w nowe konteksty. Wnioski to raczej

takie karpie statystycznego świata – trzeba traktować je z pewną ostrożnością, bo generalnie to pycha, ale miewają ości i was co prawda te ości nie zabijają, ale mnie kiedyś tak. Tak naprawdę błąd ekologiczny rozumowania, jak i wiele błędów wnioskowania, wynika trochę z naszej tendencji do ułatwiania sobie życia – jest to próba wyjaśnienia skomplikowanego zjawiska za pomocą jedynie dwóch zmiennych. Bo tak naprawdę lubimy wiedzieć i lubimy szukać wyjaśnień. Dobrym przykładem jest tu mój szef i jego reakcja na oszałamiającą anegdotkę o statystycznym lwie. No bo gdy opowiedziałam mu o tym lwie, co mówi: Rho!, Rho!, to mój szef jakoś tak za bardzo się nie śmiał, ale bardzo dużo wzdychał. Westchnął. Westchnął. Znow westchnął. *Janina...* – powiedział. Przerwał. Westchnął jeszcze raz. Spróbował znowu. *Please tell me you drink in the morning. There needs to be an explanation.*

W gruncie rzeczy wszyscy jesteśmy jak mój szef – chcemy wiedzieć, szukamy wyjaśnień i połączeń pomiędzy czynnikami. Ale z tym musimy niezwykle uważać, o czym swego czasu przypomniał nam... martwy łosoś.

Czego statystyka nauczyła się od martwego łośosia?

Zawsze, gdy myję stopy pod prysznicem, to myślę o śmierci. To skojarzenie nasuwa się samo – jak tak się mydli podeszwy odpowiedzialne za utrzymanie w pionie całego człowieka, a następnie stawia je na mokrej powierzchni, to stamtąd już tylko pół (namydlonego) kroku od śmiertelnego wypadku. Tym samym zawsze wtedy nachodzą mnie refleksje na temat kwestii ostatecznych. Na przykład tego, czy istnieje sprawiedliwość na tym świecie i tym samym wieczne potępienie czeka na tych, którzy rzucają tę prześmieszoną anegdotkę: „Ha, ha, a czy słyszałaś to powiedzenie, że istnieją trzy rodzaje kłamstw na tym świecie: kłamstwa, cholerne kłamstwa i statystyki?”¹⁰⁹ .

To znaczy umówmy się, ja mam ogromną nadzieję, że ci ludzie trafiają do piekła, skoro już do więzienia nie mogą, jak poinformował mnie pan policjant na komisariacie, gdy przyszłam złożyć zeznania. Słuchajcie, statystyki nie kłamią. Ale ludzie bardzo często kłamią na temat statystyk. Tak przynajmniej przez dłuższy czas odpowiadałam na tę zaczepkę, ale to było zanim – wraz z wszystkimi akademikami tego świata – wzięłam udział w najważniejszej lekcji statystyki we wszechświecie. A udzielił nam jej... martwy łośos.

Łosos był rybą godną – ważył niecałe 2 kilogramy i mierzył całe 45 centymetrów, co sprawiało, że paradoksalnie nie miałby wstępu na większość zjeżdżalni wodnych w aquaparkach, niemniej nie przeszkodziło mu to dostać posady na uniwersytecie w Santa Barbara. A zatrudnił go Craig M. Bennett, badacz i psycholog, który postanowił swego czasu przeprowadzić badanie, w którym wykorzystał nieinwazyjną metodę obrazowania aktywności mózgu, zwaną funkcjonalnym rezonansem magnetycznym (fMRI). Badanie to polega na tym, że człowiek wjeżdża do ogromnej tuby, jakby gigantycznej rurki z kremem (z tym że bez kremu, czyli gorzej), a w tejże maszynie odpowiednie impulsy elektromagnetyczne wchodzi w reakcje z hemoglobina i w efekcie otrzymujemy kolorową mapę aktywnych obszarów mózgu¹¹⁰ .

Bohater tego rozdziału, Craig Bennett, miał taki pomysł, by za pomocą fMRI śledzić reakcje emocjonalne swoich badanych na pokazywane im zdjęcia. Zdjęcia, które miały przedstawiać ludzi w różnych sytuacjach

społecznych. Zacząć musiał od spraw szalenie technicznych – skalibrowania maszyny fMRI. W tym celu zrobił to, co na jego miejscu zrobiłby każdy z nas – wsadził do rezonansu magnetycznego dynię.

Eksperymenty z dynią okazały się nie do końca udane. To znaczy owszem, zdjęcie wyszło ładne, do paszportu jak znalazł, niemniej Bennett nie wziął pod uwagę jednej, dość istotnej wady tego warzywa. Otóż dynia nie ma mózgu. A żeby poprawnie skalibrować maszynę do rezonansu magnetycznego, jednak niezbędny był obiekt, który ma tkanki, mięśnie, narządy wewnętrzne... Mhm, dobrze się domyślcie – chwilę później Bennett stał w kolejce w sklepie rybnym i prosił sprzedawcę o jakiś dorodny okaz. Nie był wybredny, jeśli chodzi o gatunek, najważniejsze, aby nie był to na przykład wieloryb, bo taki toby się do maszyny nie zmieścił, ani płaszczka, bo ta toby się mogła rozlewać na boki, a nikt nie ma czasu krzyczeć do pielęgniarki, żeby przytrzymała pacjenta, bo mu na podłogę płetwa zjechała.



Trafiło na łososia. W swoim artykule (absolutnie wspaniałym!) Bennett i jego ekipa badawcza bardzo rzetelnie opisali przygotowanie łososia do badania: włożyli go do maszyny, zainstalowali specjalne lustro na czole (czy łososi mają czoło?), żeby lepiej widział, kazali mu się nie ruszać (co – biorąc pod uwagę, że był martwy – wyszło mu śpiewająco), a także z uwagą obserwować pokazywane mu zdjęcia (tu obawiali się, że łososiowi może pójść trochę gorzej, bo ciągle był martwy). A potem się zdziwili, bo gdy spojrzeli na wyniki swoich badań, to dane jednoznacznie wskazywały, że łosoś jak najbardziej reagował na pokazywane zdjęcia, poszczególne części jego mózgu były aktywne. Pokazali mu zdjęcie bawiących się dzieci

– łośoś się ucieszył, pokazali klauna – łośoś się ubawił, a potem pokazali mu zdjęcie człowieka łowiącego ryby i łośoś się wkurwił.

Nie no, to ostatnie to żart. Niemniej totalnie sobie wyobrażam, że mogłoby tak być.

Wyniki analiz nie pozostawiały złudzeń: aktywne obszary łośosiego mózgu jasno wskazywały na to, że łośoś reaguje na pokazywane mu zdjęcia. Co nie byłoby znów aż tak dziwne, gdyby nie to, że – przypominam – łośoś był zimny jak ryba.

Przyznam, że chciałabym być w laboratorium Bennetta w momencie, gdy odkrył on, z jak emocjonalną rybą ma do czynienia. Ile znacie ryb, które wybuchają śmiechem nawet po tym, jak przerobiono je na paluszki? No naukowcy też znają ich niewiele.

Ale nie jest tak, że hipotezę o bogatym życiu duchowym łośosi naukowcy odrzucili od razu, niemniej postanowili też poszukać innego rozwiązania. A to ukryte było w ich obliczeniach.

Dlaczego łośoś się ucieszył?

Analiza statystyczna zastosowana w omawianym badaniu polega na wielokrotnych testach porównujących ze sobą kolejne obszary mózgu. W czasie funkcjonalnego rezonansu magnetycznego badany obszar dzielony jest na liczne trójwymiarowe piksele, tzw. woksele. Następnie każdy woksel jest porównywany z pozostałymi i na tej podstawie oznacza się go jako aktywny lub nieaktywny. W przypadku badania z łośosiem tych wokseli było aż 130 000, a tak ogromna liczba testów statystycznych przeprowadzanych symultanicznie niemal zawsze skutkuje BŁĘDEM I RODZAJU – jednym z dwóch błędów, które możemy popełnić przy testowaniu hipotez statystycznych.

Albowiem musicie wiedzieć, że każde badanie naukowe opiera się na testowaniu określonych hipotez zerowych (kierunkowych lub niekierunkowych).

PRZYKŁAD HIPOTEZY NIEKIERUNKOWEJ: ludzie karmiący się pączkami różnią się poczuciem szczęścia od tych, którzy kąpią się w sałacie.

PRZYKŁAD HIPOTEZY KIERUNKOWEJ: ludzie karmiący się pączkami są szczęśliwsi od tych, którzy kąpią się w sałacie.

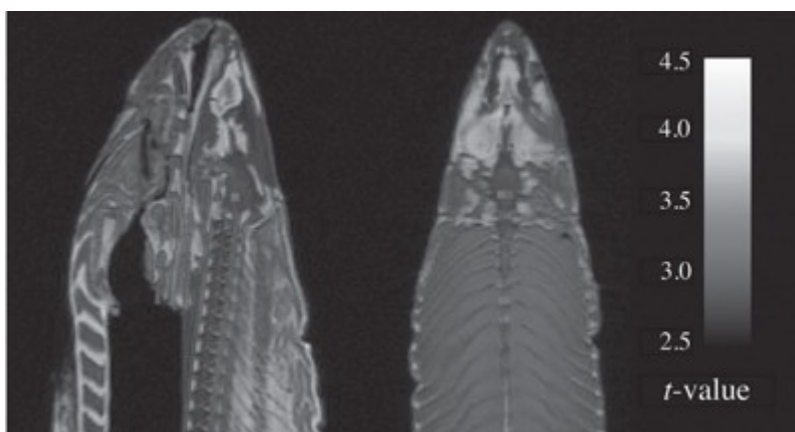
Te hipotezy następnie weryfikujemy za pomocą odpowiednich analiz statystycznych, tj. przyjmujemy je jako prawdziwe lub odrzucamy. Teraz uwaga, istnieje pewna dyskryminacja językowa dotycząca hipotez zerowych. Mówimy „odrzucać hipotezę zerową” i to samo w sobie kojarzy się źle, wie to każdy absztyfikant. Hipoteza zerowa oznacza po prostu brak zakładanego efektu. Czyli odrzucić hipotezę zerową oznacza tyle co pokazać, że „coś się w badaniu wydarzyło”. Pojawił się jakiś efekt.

Dlatego w nauce to odrzucenie hipotezy zerowej nie powinno kojarzyć się źle – weźmy przykład z testowaniem nowego leku. W takim przypadku nasza hipoteza może brzmieć tak: stan pacjentów łykających nowe tabletki będzie taki sam jak przed leczeniem. Tym samym analizy, które wskażą konieczność odrzucenia hipotezy zerowej, będą radosnym wydarzeniem, albowiem pokażą, że nasz nowy lek może działać!

Z tym testowaniem hipotez są związane dwa błędy. Ponieważ Jerzy Neyman i Egon Pearson, którzy jako pierwsi je opisali, zostali obdarzeni kreatywnością na poziomie parku wodnego Tychy¹¹¹, to owe błędy mają szalenie chwytliwe nazwy: błąd pierwszego rodzaju i błąd drugiego rodzaju.

BŁĄD I RODZAJU (*false positive*) to ten, z którym mamy do czynienia wtedy, gdy odrzucamy hipotezę zerową, która tak naprawdę nie jest fałszywa. To zupełnie taki sam błąd, jak gdybyśmy otrzymali powiadomienie o zainfekowanym pliku, z którym tak naprawdę wszystko jest w porządku¹¹². Lub informację, że jeden pączek jest gorszy od drugiego, chociaż tak naprawdę niczym się nie różnią i oba są pycha. Inny przykład: błąd I rodzaju miałby miejsce wtedy, gdyby okazało się, że według wyników testu Radek z Ursynowa (lub jakkolwiek inny mężczyzna) jest w ciąży.

BŁĄD II RODZAJU (*false negative*) to błąd odwrotny. Występuje wtedy, kiedy nie odrzucamy hipotezy zerowej, która jest w rzeczywistości fałszywa. Czyli zupełnie tak, jak gdyby test wykazał, że kobieta na tydzień przed porodem wcale nie jest w ciąży. Lub gdyby oprogramowanie nam powiedziało, że śmiało możemy otworzyć poniższy załącznik, bo jest bezpieczny, no ale okazuje się, że jednak nie.



W przypadku naszego tytułowego martwego łososia symultanicznie porównywano ze sobą ogromne ilości wokseli. Testowano tym samym hipotezę, że aktywność badanego woksela wynosi 0 (jest on więc nieaktywny). Na podstawie odpowiednich analiz ta hipoteza została odrzucona, co świadczyło o tym, że badany obszar mózgu łososia jest aktywny. Mądry komputer nawet zaznaczył badaczom te obszary łososiowego mózgu na czerwono, żeby nikt nie przegapił, że łosoś podjarał się klaunem.

Gdy dokonujemy analiz statystycznych, za każdym razem sami decydujemy, na jakie ryzyko popełnienia błędu I i II rodzaju się decydujemy. To oznacza, że im mniejsze prawdopodobieństwo popełnienia danego błędu przyjmiemy, tym większą mamy pewność, że nasze wyniki badań będą zgodne ze stanem faktycznym. Niemniej okazuje się, że jeśli nasza analiza obejmuje prawie 130 000 testów porównawczych, to ryzyko popełnienia błędu I rodzaju, a więc otrzymania wyników niezgodnych ze stanem faktycznym, jest niemal stuprocentowe. Jest to znany w statystyce **PROBLEM PORÓWNAŃ WIELOKROTNYCH**.

Istnieje wiele prostych narzędzi statystycznych, które rozwiązują ten problem, a których Bennett zapomniał zastosować (charakteryzują się tym, że mają długie nazwy: poprawka Bonferroniego, poprawka Holma-Bonferroniego czy procedura Benjaminiego-Hochberga). Myślcie o tych narzędziach jak o niezwykle skutecznych plastrach na krwawiącą ranę błędnych analiz. Serio, bo jak tylko Bennett i jego naukowa drużyna zastosowali odpowiednią korektę do swoich obliczeń, to mózg łososia natychmiast wrócił do prawidłowego, bardzo martwego stanu.

Lubię ten przykład, bo nawet jeśli ktoś nie interesuje się technikaliami, to powinien zrozumieć, że możemy mieć dostęp do najlepszej jakości danych, ogromnej mocy obliczeniowej i wspaniałej grupy badawczej, ale

w gruncie rzeczy badacze wciąż są ludźmi i czasem popełniają błędy. Co nie znaczy, że wszystko, co robią, nadaje się do kosza i że nie można mieć do nich krzty zaufania. Tego nauczyli nas martwy łosoś, a także pewien człowiek szalenie zaniepokojony losem dzięciołów i ich podatnością na migreny.

Dlaczego od stukania w drzewo dzięcioła nie boli głowa?

Przypadek martwego łososa, który wzruszał się na zdjęciach, to jeden z najbardziej znanych przykładów błędów badaczy wynikających z nieprawidłowych obliczeń. Craig Bennett wraz ze swoją drużyną dorobił się nawet za swoje odkrycie Nagrody Ig Nobla. Choć łososiowi przyznano ją *post mortem*, to rodzina i sąsiedzi i tak byli z niego dumni.

Nagrody Ig Nobla to w ogóle fantastyczne wydarzenie, choć często wyśmiewane i traktowane niezbyt serio. Prawdopodobnie do poprawy jego wizerunku nie przyczynił się Ivan Schwab, który w 2006 roku otrzymał tę nagrodę za badanie, mające wyjaśnić, dlaczego dzięcioła nie boli głowa od ciągłego uderzania w drzewo¹¹³, a swoją statuetkę odebrał w przebraniu tegoż ptaka.

Niemniej tak naprawdę rozdanie Nagród Ig Nobla to święto nauki niestandardowej (nie mylić z alternatywną!) i kreatywnego myślenia. Odbywa się zresztą na Uniwersytecie w Harvardzie, z którego raczej się nie żartuje. No chyba że jest się z Oksfordu, to wtedy można. A mottem tego wydarzenia jest zdanie: „nagroda za odkrycia naukowe, które powodują śmiech, a następnie zmuszają do myślenia”¹¹⁴.

Podobnie było z martwym łososiem – badanie fMRI ryby, które na początku potraktowane zostało jako żart, tak naprawdę doprowadziło do bardzo konkretnych zmian w metodologii badań. Zwróciło uwagę na konieczność stosowania odpowiednich korekt statystycznych przy tej konkretnej metodzie analizy¹¹⁵. Dziś prawie wszyscy pamiętamy o tym, by – tam gdzie to konieczne – stosować poprawki dla porównań wielokrotnych, niemniej niewielu pamięta, że tę wiedzę zawdzięczamy pewnemu miłemu rybnemu mieszkańcowi Atlantyku¹¹⁶.

Ostrzeżenia o liczbach i obliczeniach, które mogą wprowadzać w błąd, przeczytamy zarówno w artykułach medycznych (Amin et al. 2014), jak i tych z dziedziny biologii (Zar 1968) czy fizyki (Chen et al. 2008). Nie świadczy to o słabości tej metody, wszak jak mawiał bohater filmu *Pół*

żartem, pół serio: „nikt nie jest idealny”¹¹⁷ (choć najpewniej nie miał na myśli statystyki). Taka jest rola nauki, że ciągle dowiadujemy się czegoś nowego i czasem przychodzi nam weryfikować własne założenia czy teorie. Co oczywiście nie zawsze jest proste, jestem w stanie sobie wyobrazić minę Ptolemeusza, gdy usłyszał, że ktoś właśnie obalił jego koncepcję heliocentrycznej budowy świata. Dość powiedzieć, że Kopernika nigdy na wigilię nie zaprosił.



Dla mnie naukowy sposób myślenia o świecie oznacza, że nigdy nie jest się niczego stuprocentowo pewnym. Nauka zaczyna się od wątpliwości i one tak naprawdę nigdy się nie kończą. To dobrze, że możemy otwarcie dyskutować o tym, że nasze metody mogą być obarczone błędem, a także szukać rozwiązań tego problemu. To dobrze, że wciąż weryfikujemy wszystko, co wiemy, i nigdy nie jesteśmy pewni. To nawyk, który powinniśmy w sobie wyrobić i pielęgnować go czule niczym rzeżuchę, którą w podstawówce własnoręcznie zasadziliśmy na wacie. Zwłaszcza że gdy stracimy czujność, to możemy wpaść w pułapkę szarlatanów, takich trochę współczesnych alchemików – takich, którzy już nie zamieniają ołowiu w złoto, ale zamieniają czysty, rodzicielski strach w wiadra złotych monet. Dlaczego więc w ogóle im ufamy?

Dlaczego ufamy szarlatanom?

Czasem warto zwrócić się o pomoc do ludzi mądrzejszych od siebie. Do mnie ta prawda dochodzi zwłaszcza w okolicy czerwca, bo czerwiec to taki miły czas w roku, kiedy to staję się najbardziej pożądaną gazelą w uniwersyteckiej dżungli, szafranem studenckiego świata i najpiękniej wyczesanym owczarkiem collie wśród wykładowców akademickich, co oznacza mniej więcej tyle, że wraz ze zbliżającym się terminem oddawania prac dyplomowych w sposób nagły wzrasta studencka miłość do instytucji liczenia, a raczej do ludzi, którzy posiadli tę magiczną wiedzę, że logarytm naturalny to wcale nie jest gatunek sera. No a co mam wam powiedzieć – akurat liczby, zupełnie odwrotnie niż własny apetyt, to ja poskrwiam doskonale.

No i ja sobie jestem czasem taką szalenie pożądaną gazelą, ja sobie tak kopytkuję co rano do biura, ale to kopytkowanie ma w sobie często pewien pierwiastek dramatyzmu, albowiem bardzo się staram robić to w sposób niezauważalny dla otoczenia, tak by po drodze nie zostać trafiona pociskiem studenckich pytań, a to jest szalenie trudne, bo jestem człowiekiem o żenującej zdolności koncentracji, więc jak tak sobie udaję przebiegłą gazelę, to zawsze w pewnym momencie tracę czujność, bo akurat się rozkojarzę myśleniem o tym, że w sumie to takie gazy mają super, bo nie muszą golić nóg.

A ostatnio to już w ogóle klęska, bo tak doskonale szło mi to kopytkowanie po korytarzu, byłam ledwie trzy kroki od bezpiecznej przestrzeni własnego biura, a wtedy ciach, zobaczyłam za oknem dwie mewy i to było silniejsze ode mnie, no nikt by się nie powstrzymał, natychmiast zaczęłam się zastanawiać nad tym, czy mewy mogą być homoseksualne. No bo pingwiny na przykład mogą i delfiny też, a z delfinów to w ogóle są takie gagatki, że niektóre się nawet zakochują w żółwiach morskich albo węgorzach (serioszka!), a o mewach to ja akurat nic nie wiedziałam, więc się zainteresowałam, i akurat jak się nad tym zastanawiałam, to ciach, wykorzystując moje chwilowe rozkojarzenie, dopadł mnie student i powiedział: „Cześć, Janina, co robisz?“, a ja powiedziałam: „Cześć, Conor, co wiesz o mewach?“.

Ku mojemu rozczarowaniu Conor niewiele wiedział o mewach, niemniej miał bardzo dużo ważnych przemyśleń na temat życia i śmierci, ze

szczególnym wskazaniem na to drugie, co z kolei miało ścisły związek z bliskim terminem oddania jego pracy licencjackiej. I on mnie spytał, czy mogłabym mu pomóc w jednym modelu regresji logistycznej, a ja przewróciłam oczami i westchnęłam, że Conor, Conor, Conor, takie podejście do obowiązków jest naprawdę niepoważne, przecież jesteśmy dorosłymi ludźmi, organizacja pracy jeszcze nigdy nikogo nie zabiła, i jeszcze powiedziałam wiele innych rzeczy, które bardzo często powtarza mi mój szef. Ale musicie wiedzieć, że tak naprawdę to trochę udawałam z tym zniecierpliwieniem, bo regresja logistyczna jest moją ulubioną z wszystkich regresji, więc to było trochę tak, jakby ktoś właśnie rzucił mi w twarz najpiękniejszym w świecie schabowym i jeszcze dodał wiadro mizerii.

Więc idziemy to policzyć i jak tak idziemy, to ja jestem najbardziej zadowolonym człowiekiem świata, za to nasz kolega Conor to tak nie za bardzo, on na buzi to ma taki odwrócony rogalik, bardzo jest nieszczęśliwy i mówi, że on naprawdę nie rozumie, do czego mu się to wszystko kiedykolwiek w życiu przyda, a ja mu na to mówię, że akurat liczenie to mu się może w życiu okazać pomocne, na przykład wtedy, gdy pójdzie kiedyś z dziećmi do supermarketu, a potem źle je policzy w drodze powrotnej i żona mu w domu powie, że w sumie całkiem spoko, że wrócił z trójką dzieci, ale czemu dwójki brakuje. I wtedy Conor, może być ci bardzo trudno wytłumaczyć żonie, czemu się nie doliczyłeś własnych dzieci, bo o ile jedno zgubione dziecko można jeszcze potraktować jako błąd pomiaru, to dwójka już najpewniej nie przejdzie.

A poza tym Conor – mówię mu – uważam, że człowiek musi się w życiu uczyć wielu znacznie bardziej zbędnych rzeczy niż matematyka, na przykład ja całą podstawówkę byłam traumatyzowana nauką fikołków i do dziś nie wiem, do czego mi te fikołki były potrzebne, a wtedy on patrzy na mnie, patrzy, i to jest taki wzrok totalnego niezrozumienia, znaczy wiecie – ja go pytam, do czego w ogóle przydaje się w życiu umiejętność robienia fikołków, a on patrzy na mnie jak na człowieka idiotę i mówi: Nooo... – tłumaczy mi takim tonem, jakby przekazywał najbanalniejszą prawdę tego świata – żeby bawić się z kolegami w ludzkie kręgle.

Słuchajcie, warto mieć w życiu swojego Conora.

Taką osobę, z którą można dzielić się zarówno umiejętnościami, jak i frytkami w McDonalddie, co jest – umówmy się – najwyższą formą miłości i oddania, plasującą się nieznacznie wyżej niż oddanie komuś nerki. Taką

osobę, której my nakreśliliśmy meandry regresji logistycznej, a ona w zamian udzieli nam odpowiedzi na pytania egzystencjalne dotyczące użyteczności fikołków.

Ja mam jeszcze jedną taką osobę – jest to absolutnie najmądrzejszy człowiek, jakiego znam, i to nie tylko dlatego, że umie składać prześcieradło z gumką (!!!). Oddałabym mu frytkę, nerkę i w ludzkie kręgle też bym z nim zagrała. Proszę państwa, przed państwem człowiek, który wie absolutnie wszystko o błędach poznawczych – Piotr Bucki.

Nie było nikogo innego, kogo mogłabym poprosić o napisanie rozdziału dotyczącego mechanizmów, które powodują, że wierzymy szarlatanom, wsadzamy dzieci na trzy zdrowaśki do pieca i wydajemy wiadra złotych monet na butelkę wody, która ponoć ma lepszą pamięć niż Jadwiga z Supraśla, która to wyszła kiedyś do sklepu bez listy, mój Boże, co za heroizm, i ona wyszła po jajka, chleb i kostkę masła, a wróciła z hulajnogą i trzema kilogramami cementu. Piotr wam to wszystko wyjaśni znacznie lepiej ode mnie, więc zapraszam do czytania. To będzie intelektualna przygoda jeszcze lepsza niż gra w ludzkie kręgle.

Belle gibson i nadzieja. Piotr Bucki

5 marca 2015 roku w Australii Belle Gibson uczestniczyła w pogrzebie przyjaciółki, Jessiki Ainscought. Kobiety sporo łączyło i dzieliło jednocześnie. Belle Gibson dzięki cudownej diecie, naturoterapii i organicznemu jedzeniu pozbyła się raka, którego tak po prawdzie nigdy nie miała. Jessica stosowała te same metody, jednak naprawdę miała raka.

Australijczycy to społeczeństwo bogate, w miarę bezpieczne i ze wszech miar szczęśliwe¹¹⁸. Bardziej niż ataku nuklearnego boją się glutenu, pestycydów i dymu z papierosów. Furorę robią tam żywność organiczna, urządzenia do alkalizowania wody oraz wszelkie terapie alternatywne z etykietą New Age.

Australia to trochę taki ekskluzywny raj dla szarlatanów. Nie spotkasz tam znachorów, którzy radzą wsadzić dzieciaka na trzy zdrowaśki do pieca, lub szeptuch z Podlasia, które będą zaklinały kurze jajka, by wypędzić demony. Zobaczysz za to pięknych ludzi w laboratoryjnych fartuchach, którzy zapraszają przyjaznym gestem i uśmiechem białych licówek Da Vinci do gabinetów na konsultacje.

W tym właśnie świecie interes życia – do czasu – robiła Belle Gibson.

Gibson chętnie opowiadała o swojej dramatycznej walce z rakiem. W mediach, w przedmowie do swojej książki *The Whole Pantry*. Utrzymywała, że rak zaatakował z całą siłą swej złośliwości mózg, krew, śledzionę, wątrobę, płuca i nerki. Odpowiedzialnością za tego złośliwego wielonarządowego potwora obarczyła Gardasil – szczepionkę zapobiegającą rakowi szyjki macicy. I gdy nie było już nadziei, a tradycyjna medycyna rozłożyła z bezradności ręce, Gibson zaczęła sama szukać lekarstwa.

Gdzie?

W domowej spizarce. Przeszła na stricte wegańską dietę bazującą na skrupulatnie dobranych organicznych składnikach. Z pasją włączała do posiłków tzw. superfoods¹¹⁹. Eksperymentowała z radykalną terapią (protokołem Gersona), która ma rzekomo leczyć raka¹²⁰. I wtedy stał się cud. Rak zniknął.

Gibson swój sposób na leczenie opisała w książce. Stworzyła też aplikację *The Whole Pantry*, która pomagała w zdrowych wyborach i wspierała (rzekome) leczenie. Wygrana z chorobą przełożyła się na spory sukces. Medialny i finansowy. W końcu kto by nie chciał przeczytać bajki o dzielnej księżniczce, która odrzuciła patriarchalne ramię konwencjonalnej medycyny i sama zgładziła smoka.

I taką właśnie opowieść zaserwowała czytelnikom australijska edycja magazynu „Elle”. Jakiś czas po opublikowaniu materiału do redakcji przyszedł mail. Ktoś napisał: „Nie promujcie tej kobiety. Ona kłamie. Nigdy nie miała raka”. Dziennikarze nie dowierzali. Postanowili jednak to sprawdzić. Okazało się, że Gibson rzeczywiście wymyśliła sobie chorobę, żeby uwiarygodnić lekarstwo. By uprawdopodobnić cud.

Kłamstwo to bardzo skomplikowany mechanizm. Równie skomplikowane mechanizmy stoją za decyzjami tych, którzy w te kłamstwa uwierzyli. Ba, którzy nawet po demaskacji i procesie sądowym, w wyniku którego Gibson skazano za działanie na szkodę klientów, nadal uważają, że jej „ukryte” terapie działają.

Na początek – ważna jest NADZIEJA. Umiera ponoć ostatnia. Dlatego nie ma nic dziwnego w tym, że ktoś, komu medycyna tradycyjna nadzieję raczej odbiera lub przedstawia w formie statystycznego prawdopodobieństwa, może wybrać narrację dającą nadzieję na wyzdrowienie. Nawet w formie anegdotycznego dowodu (pamiętacie? To

ten zły brat bliźniak metody naukowej, o którym pisaliśmy więcej w rozdziale *Czy naukowcy są jak żony i nigdy się nie myślą?*).

W tym przypadku nadzieję podano w formie HISTORII. I właśnie w samej formie narracji tkwi kolejna pułapka, w którą mogli wpaść ci, którzy uwierzyli Gibson. Prawdziwe informacje podane w sposób sfabularyzowany stają się dla naszego mózgu prawdziwsze – bardziej wiarygodne (Haase et al. 2015). Niestety tak samo działa to w przypadku nieprawdziwych informacji. Innymi słowy, kategoria obiektywnej prawdy jest raczej obca naszemu intuicyjnemu, szybkiemu oglądowi świata. Mózg lubi być w stanie ŁATWOŚCI POZNAWCZEJ, czyli w takim, w którym wszystko rozumie, a poszczególne elementy narracji plotą się w jeden piękny, logiczny, rumiany warkocz chałki o znajomym smaku. Ten smak to podobieństwo i struktura. Im bardziej coś swojskie, tym łatwiej nam uwierzyć, że wydarzyło się naprawdę. To wina HEURYSTYKI DOSTĘPNOŚCI i REPREZENTATYWNOŚCI. Im bardziej jesteśmy sobie w stanie przypomnieć historie o innych przypadkach cudownego uzdrowienia, tym łatwiej jest nam uwierzyć w kolejne.

I tak było w przypadku Gibson. Jej historia była bardzo podobna do bajek, legend i filmowych scenariuszy, w których bohater, wydawałoby się bez szans, pokonuje smoka, potwora czy innego lorda Vadera, bo odkrywa magiczny napój, znajduje ukrytą moc lub poznaje stare zaklęcie. I choć na poziomie racjonalnym trudno uwierzyć, że tak to działa, naprawdę jesteśmy bardziej skłonni wierzyć w historie, które przypominają nam bajki z dzieciństwa. Podobieństwo struktur, zdarzeń i narracji wpływa na nasze nieświadomione procesy przetwarzania treści.

Wpływa tak mocno, że niektóre historie – choć fałszywe i szkodliwe – rozprzestrzeniają się jak wirusy. Zwłaszcza, gdy spełniają pewne warunki, które określił w książce *Contagious* (polskie wydanie – *Efekt wirusowy w biznesie*) Jonathan Berger (2013). Profesor Uniwersytetu Stanforda wyłuskał wspólne dla wszystkich wirusowych (prawdziwych i nieprawdziwych) historii cechy, które wzmacniały znacząco tzw. transmisję społeczną¹²¹. Na to, czy historia stanie się wirusowa (jak choćby w przypadku historii Gibson), wpływają następujące jej cechy:

- WALUTA SPOŁECZNA (ang. *Social currency*¹²²), czyli to, czy informacja jest wartościowa dla odbiorcy. Dobry dowcip, historia dająca nadzieję czy soczysta plotka dają mocną walutę społeczną. Suchar, niepozostawiająca złudzeń lub zwyczajnie nudna historia czy też plotka,

o której wszyscy już słyszeli, tej wartości nie mają. Trzeba jednak pamiętać, że sama waluta społeczna może dość szybko ulegać inflacji. Jej wartość zależy od wysycenia społeczności daną informacją. Kiedy ludzie osłuchają się już z dowcipem czy – choćby najlepszą – historią, tracą nimi zainteresowanie.

- **PRZYPOMINAJKA** (ang. *Trigger*), czyli to, czy w danej informacji zaszyte są jakieś informacje, które pozwalają na łatwe jej odtworzenie. Na przykład historię Belle Gibson ludzie przywoływali sobie za każdym razem, gdy słyszeli o kimś, kto choruje na raka, albo za każdym razem, gdy słyszeli o protokole Gersona. Informację z naszej pamięci przywołują określone bodźce. Czasem słowa, czasem zapachy, a czasem widok jakiegoś obiektu.

- **EMOCJE** (ang. *Emotion*), czyli to, czy i jakie historia lub informacja budzi emocje. Jeśli nie budzi ich wcale, to szybko popadnie w zapomnienie. Emocje nie tylko kierunkują uwagę. Wpływają też na zapamiętywanie i, pośrednio, na uwiarygodnienie informacji. Oraz na transmisję społeczną (patrz przypis na str. 265).

- **PUBLICZNOŚĆ** (ang. *Public*), czyli to, czy informacja, którą przekazemy odbiorcy, wzmocni naszą pozycję w grupie. Informacje przekazujemy innym, kierując się przy tym dwoma motywacjami – altruistyczną i narcystyczną. Podajemy historie dalej, bo chcemy obdarować nimi innych. Ale chcemy też, żeby inni postrzegali nas jako inteligentniejszych, zabawniejszych, lepiej poinformowanych.

- **WARTOŚĆ PRAKTYCZNA** (ang. *Practical value*), czyli to, czy informacja, którą przekazujemy, ma dla odbiorcy wartość praktyczną. Sama wartość praktyczna jest oczywiście subiektywna. Nie musi być też rozumiana jako sformułowana klarownie instrukcja, jak zrobić chłopaka z kartonu po mleku czy jak upleść koszyk z kabanosów. Może też wynikać z tego, że dowiadujemy się, że istnieje szansa wygrania z rakiem i dieta, która nam w tym pomoże.

- **SEKWENCJA WYDARZEŃ** (ang. *Story*), czyli to, czy informacja ma formę historii, anegdoty lub innej chronologicznej sekwencji zdarzeń. Wynika to z łatwości poznawczej. Historie ułatwiają zrozumienie i są łatwe do przekazywania dalej.

Wyznawcy teorii Gersona, Belle Gibson i pewnego inżyniera maszyn budowlanych wierzą nie tylko dlatego, że karmią się nadzieją. Popędzają

też pewien błąd poznawczy, który nazywamy BŁĘDEM NATURALISTYCZNYM lub odwołaniem do natury. Odwołanie do natury to argument lub retoryczna zagrywka, w której zjawisko lub rzecz opisuje się jako pożądane, ponieważ jest naturalne. Analogicznie: niepożądane jest to, co nienaturalne. Przy czym naturalne jest wszystko to, co w mniemaniu argumentującego pochodzi od natury. A nienaturalne (*ergo* złe) jest to, co stworzył człowiek. Argumentu tego używa się często w dyskusjach na temat homoseksualizmu, ekologii, medycyny, wegetarianizmu, feminizmu czy fizyki.

Jest to argument o tyle chybiony, że nie bierze pod uwagę tego, że wiele przywilejów, którymi cieszymy się współcześnie, wynika właśnie z „nienaturalnej” działalności człowieka. Naturalne jest to, że prawie każde ciężkie zakażenie bakteryjne czy wirusowe kończyło się śmiercią. Naturalne jest to, że po ciężkich złamaniach w najlepszym wypadku czekało nas trwałe kalectwo. Naturalna byłaby śmierć przedwcześnie urodzonych dzieci. Wreszcie całkowicie naturalna jest poligamia i częste, zupełnie nieposkromione zapładnianie przez samca (mężczyznę) jak największej liczby samic (kobiet).

W przypadku historii Gibson mieliśmy do czynienia z typowym dla „naturalistów” polem bitwy. Medycyna konwencjonalna kontra medycyna „naturalna”. W tej bitwie w (nieprawdziwej) opowieści Gibson wygrała zdrowa natura. Przegrała zaś zła medycyna. Odwołanie do natury jest częstsze w czasach, w których technologia stała się tak skomplikowana, że – dla większości z nas – jest zupełnie niezrozumiała. Idealizowanie natury jest więc mechanizmem obronnym. Robimy to również z dzieciństwem lub starymi dobrymi czasami. Nic tak nie sprzyja gloryfikowaniu natury jak cywilizacja, w której czujemy się od tej natury oderwani. Medycyna w takim świecie jawi nam się nie tylko jako coś, co jest nienaturalne. To coś niezrozumiałego, obcego i na dodatek niedającego stuprocentowych gwarancji. Nic dziwnego, że wolimy „proste recepty”, które – jak nam się zdaje – działają. Dlaczego? Bo czasem ulegamy efektowi placebo. Czasem zaś popełniamy błąd. Nosi on nazwę *post hoc ergo propter hoc* (z łac. po tym, więc skutek tego). To błąd logiczny, który występuje wtedy, gdy ktoś wyciąga wniosek o relacjach przyczynowo-skutkowych ze zwykłego następstwa dwóch zdarzeń w czasie (szerzej opisane jest to w rozdziale *Ile trzeba zjeść czekolady, żeby dostać Nobla?*).

Błąd ten najczęściej wiąże się z emocjami – a konkretnie z niepewnością. Jako ludzie jesteśmy maszynami do postrzegania wzorców. W naturze i w życiu społecznym. Nie znosimy niepewności. Wolimy założyć, że A powoduje B, ponieważ najpierw było A, a potem zdarzyło się B. Ta obserwacja bywa nawet trafna. No właśnie. Bywa, a nie jest.

Wierzymy, bo chcemy wierzyć. Bo chcemy mieć nadzieję w świecie, gdzie tej nadziei brakuje. Jeśli ktoś nam ją jeszcze wzmocni krzepiącą historią o pięknej dziewczynie, która wygrała z rakiem, to ulegamy niebezpiecznej iluzji. Groźnej dla naszego zdrowia.

Przyjaciółka Gibson zmarła. Chorowała (naprawdę) na rzadką postać raka. Ratunkiem byłaby amputacja ręki. Jednak Ainscought odrzuciła takie leczenie na rzecz kuracji „naturalnej”. Nie wyleczyła się. Zmarła też jej matka, która chorowała na występujący stosunkowo często nowotwór piersi, który zresztą został w jej przypadku bardzo wczesnie wykryty. To z kolei dawało bardzo duże szanse na całkowite wyleczenie. Jednak matka również uległa narracji o „naturalnych ukrytych terapiach”.

To po części historia większości z nas. Wierzymy, bo bardzo chcemy wierzyć. I nie chcemy wiedzieć, że mamy 65% szans na przeżycie czterech lat. Wolimy wierzyć, że w 100% wyleczą nas przecierowe soki, alkaliczna woda czy sok z kapusty.

Nie wyleczą.

Oczywiście nie zwracam się tu przeciw zdrowej diecie. W leczeniu onkologicznym żywienie jest niezwykle ważne. Co więcej, mamy coraz więcej dowodów na to, że dieta przyczynia się do ekspresji genów. I może przyczyniać się do zwiększonej zachorowalności na nowotwory. Chodzi o to, żebyście chcieli widzieć świat takim, jaki jest. A nie takim, jaki chcielibyśmy, żeby był. Bo wiara przenosi góry jedynie metaforycznie. Metaforycznie również czyni cuda. Medycyna daje realne szanse. Nie złudne nadzieje.

Inteligentni bałaganiarze i horoskopy

Jakiś czas temu wielu moich znajomych masowo zaczęło udostępniać artykuły z portalu Business Insider. Taki, którego w zasadzie nie ma sensu czytać, bo nagłówek zawiera wszystko, co trzeba wiedzieć: *Inteligentni ludzie są bałaganiarzami, późno chodzą spać i więcej przeklinają*. Jak można się domyślić, znajomi, którzy to udostępniali, nie chodzą spać z kurami

i nie przeszliby testu białej rękawiczki Perfekcyjnej Pani Domu. Obficie też kraszą język naszym narodowym przekleństwem, które wykorzystuje piękne brzmienie rozedrganej głoski r.

Dlaczego ów tekst zyskał taką popularność? Bo jego nagłówek tak bardzo oddawał całą „prawdę” o nich. Nie musieli nawet czytać tekstu. Wystarczyło, że znaleźli coś, co było zgodne z ich przekonaniem, motywacjami i światopoglądem.

To tak zwany BŁĄD POTWIERDZENIA (błąd confirmacji) i wpływ EFEKTU FORERA (Forer 1949), znanego też jako efekt Barnuma.

Błąd potwierdzenia to przywiązywanie wagi do argumentów, danych lub zdarzeń, które są zgodne z naszym punktem widzenia i oglądem rzeczywistości. To szukanie tego, co potwierdza naszą tezę. Przy jednoczesnym ignorowaniu tego, co tej tezie przeczy.

W 1948 psycholog Bertram R. Forer dał studentom do wypełnienia test osobowości. Następnie przedstawił im analizę ich osobowości rzekomo bazującą na wynikach tego testu. Studenci mieli ocenić trafność tej analizy w skali od 0 (bardzo słaba) do 5 (znakomita). Średnia ocen wyniosła 4,26. Czyli całkiem nieźle.

Później Forer ujawnił, że w rzeczywistości wszyscy studenci otrzymali tę samą analizę. Bez względu na udzielane odpowiedzi uczestnicy badania przeczytali taką oto „prawdę” o sobie:

Masz potrzebę, by ludzie cię lubili i podziwiali, jednak jesteś osobą krytyczną wobec siebie. Masz pewne wady osobowości, ale potrafisz je kompensować tym, co jest w tobie dobre. Masz duże możliwości, które wciąż pozostają niewykorzystane. O ile na zewnątrz możesz wyglądać na osobę zdyscyplinowaną i opanowaną, wewnątrz często trapi cię niepewność i martwisz się o wiele spraw.

Niekiedy masz poważne wątpliwości, czy twoja decyzja była dobra, albo czy twoje czyny były właściwe. Lubisz pewną ilość zmian i różnorodności, a kiedy osaczają cię ograniczenia, odczuwasz niezadowolenie. Cenisz sobie własną niezależność myślenia i nie przyjmujesz cudzych twierdzeń bez przekonujących dowodów.

Życie nauczyło cię, aby nie przesadzać ze szczerością, kiedy się przed kimś otwierasz. Czasem bywasz osobą otwartą na ludzi, przystępną

i towarzyską, ale innym razem zamkniętą, ostrożną i zdystansowaną. Niektóre z twoich marzeń wydają się nierealistyczne.

Trudno się nie zgodzić. Gdybym dostał taką analizę i miał ocenić jej trafność, to długo bym się nie wahał. Trafna. DLA KAŻDEGO. Lub dla prawie każdego. Dlaczego? Bo zawiera tak zwane stwierdzenia Barnuma.

To stwierdzenia zawierające opis naszej osobowości (lub odnoszące się do naszej przyszłości), które są na tyle ogólne, że praktycznie każdy może odnieść je do siebie. Na przykład takie: „jesteś w gruncie rzeczy dobry i wrażliwy, choć ludzie nie zawsze rozumieją twoje intencje”. Podobnie jak stwierdzenia zaserwowane studentom przez Forera: „masz duże możliwości, które wciąż pozostają niewykorzystane, lub czasem bywasz osobą otwartą na ludzi, przystępną i towarzyską, ale innym razem zamkniętą, ostrożną i zdystansowaną”.

No dobra, ale dlaczego te stwierdzenia nazywamy stwierdzeniami Barnuma? Przecież to Forer jako pierwszy badał zjawisko, które potem nazwano efektem horoskopowym.

Dlatego że – mimo iż Barnum nie był psychologiem – to on jako pierwszy przyjrzał się temu zjawisku. Phineas Taylor Barnum to amerykański przedsiębiorca z XIX wieku. Działał w branży cyrkowej. Takiej w ówczesnym stylu: kobiety z brodami, siłacze i menażeria bizarności rodem z filmu *Freaks* Teda Browninga z 1932 roku (totalnie polecam). No i obowiązkowo wróżki.

Oprócz tego, że był sprawnym (a co za tym idzie – bogatym) przedsiębiorcą, Barnum był również dociekliwy. Bardzo intrygowało go to, dlaczego ludzie są gotowi płacić wróżkom i wychodzą od nich zadowoleni, mimo że te powtarzają wszystkim właściwie to samo (a miał okazję to sprawdzić). Co więcej, każdy z klientów był przekonany, że wróżba dotyczyła tylko jego i bardzo dobrze go opisywała. I tak właśnie Barnum odkrył swoje stwierdzenia. Specjalizują się w nich wróżki, chiromanci i twórcy niektórych quizów na Facebooku. Jak również pseudonaukowcy i bieda-coache, którzy na podstawie przedziwnych paranaukowych testów z przedrostkiem neuro- serwują precyzyjną „wiedzę o Tobie”.

Stosując zasłonę dymną z technologii i przedrostków neuro-, wykorzystują kolejny błąd w interpretacji rzeczywistości. Otóż okazuje się, że jeśli dodamy do nazwy metody mającej pomóc ustalić czy zdiagnozować naszą osobowość ten przedrostek, to takie wyjaśnienie lub metoda

wydadzą nam się bardziej wiarygodne (Weisberg et al. 2008). Bez względu na to, czy na taką wiarygodność rzeczywiście zasługują. Nie mówię tu oczywiście o neurologii czy neurobiologii, bo to po prostu nazwy dziedzin nauki czy medycyny. Mam na myśli raczej „neurolingwistyczne testowanie koloru mózgu”. Czy też „neurolingwistyczne metody wzmacniania połączeń między lewą a prawą półkula”. Czyli metody, których nikt nie zweryfikował. A które brzmią mądrze tylko dlatego, że zawierają wspomniany przedrostek.

Szarlatanom ufamy, bo są pewni tego, co mówią. Pewnie też mówią. A to sprawia, że ulegamy kolejnemu złudzeniu. EFEKTOWI AUTORYTETU.

Doskonale pokazuje to przykład Elisabeth Holmes, bohaterki jednego z większych skandali w Dolinie Krzemowej. Od 2004 roku oszukiwała inwestorów, opinię publiczną, partnerów biznesowych i media. Twierdziła, że skonstruowany przez nią niewielki aparat o nazwie Edison będzie w stanie wykonywać ponad 270 testów diagnostycznych z kilku kropli krwi kapilarnej. Zainteresowanym polecam doskonały dokument o tej historii *Wynalazczyni: Dolina Krzemowa w kropli krwi* (2019) dostępny na HBO.

Proste gwarancje i skomplikowane dywagacje

W 1979 roku wspomniani już wcześniej Kahneman i Tversky przeprowadzili pewne badanie. Jego uczestnicy otrzymali na wstępie informację, że kraj przygotowuje się do wybuchu epidemii nieznanej wcześniej azjatyckiej choroby, która zdaniem specjalistów może doprowadzić do śmierci 600 osób. Na całe szczęście dwa zespoły naukowców pracują nad dwoma wariantami programów ratunkowych. To uczestnicy muszą wybrać, który z nich zastosują władze.

Pierwszy wariant badania wyglądał następująco: eksperci proponują dwa programy ratunkowe – A oraz B. Skutki programów to:

- PROGRAM A: zostanie uratowanych 200 osób;
- PROGRAM B: z prawdopodobieństwem $1/3$ uda się uratować wszystkich zagrożonych (600 osób), ale z prawdopodobieństwem $2/3$ nikt nie zostanie uratowany.

Jak widać, pierwszy program był przedstawiony jako opcja pewna, drugi zaś jako opcja ryzykowna. Co ciekawe, obie opcje mają tę samą wartość oczekiwaną. Są po prostu takie same.

Jednak przy takim przedstawieniu opcji większość badanych (ponad 70%) wybrała program A.

To jednak nie koniec. Badacze przygotowali drugi wariant badania. Ten wyglądał następująco:

Ekspertcy proponują dwa programy ratunkowe – C oraz D, których skutki to:

- PROGRAM C: umrze 400 osób;
- PROGRAM D: z prawdopodobieństwem $1/3$ nikt nie umrze, ale z prawdopodobieństwem $2/3$ umrą wszystkie zagrożone osoby.

Analogicznie jak w wariacie pierwszym zadaniem osób badanych było wybranie jednego spośród dwóch powyższych programów ratunkowych – C albo D. Tutaj także pierwszy program został przedstawiony jako opcja pewna, drugi zaś jako opcja ryzykowna, obie o tej samej wartości oczekiwanej. Okazuje się, że w tym wariacie większość badanych (niemal 80%) wybiera program D.

Wystarczy poświęcić chwilę na porównanie wersji A i B oraz C i D, by zauważyć, że opcje A i C oraz B i D w każdym z wariantów problemu są identyczne, choć przedstawione w odmienny sposób. Różnica polega na tym, że w pierwszym wariacie zostały uwypuklone zyski (ratowanie życia), a w drugim – straty (utrata życia). Zatem w pierwszym wariacie eksperymenci badani dokonywali wyboru między opcją pewnego zysku (program A) a opcją niepewnego zysku (program B), w drugim wariacie zaś – między opcją pewnej straty (program C) a opcją niepewnej straty (program D).

Powyższe badanie, wykonane i opisane przez Tversky'ego i Kahnemana (1979), noszące nazwę problemu choroby azjatyckiej (*Asian Disease Problem*), pokazuje, że wybór między ryzykowną a pewną opcją o tej samej wartości oczekiwanej silnie zależy od tego, czy zostały one przedstawione w kategoriach zysku, czy w kategoriach straty. W przypadku uwypuklenia zysku badani wykazują awersję do ryzyka. W przypadku uwypuklenia straty badani wykazują do niego skłonność.

Pomyślcie teraz, jak to się odnosi do wszystkich tych sytuacji, gdy biedacoach, znachor, wróżka, pseudonaukowiec lub innej maści szarlatan pokazuje wam opcję, która w przeciwieństwie do nauki, medycyny czy statystyki, proponujących rozwiązania „ryzykowne”, jest „pewna”. Tamte są ryzykowne, bo operują prawdopodobieństwem. Ryzykowne, bo pełne

niuansów. Ryzykowne, bo naukowe podejście charakteryzuje się tym, że mówi o obecnym stanie wiedzy, a nie twierdzi, że coś jest pewne na 100%.

Wolimy proste, nieskuteczne recepty. Wszyscy. Ja pewnie też. Jediną drogą jest świadomość i polubienie nauki – tej pokornej drogi, o której Carl Sagan mówi, że „pokazuje nam świat takim, jakim jest, a nie takim, jakim chcielibyśmy, żeby był”.

Wspomniany wcześniej Kahneman jest autorem jednego z moich ulubionych cytatów. Ten noblista powiedział kiedyś, że nie tylko jesteśmy ślepi na błędy, ale też: „Jesteśmy ślepi na własną ślepotę. Niewiele wiemy o tym, jak mało wiemy. Nie mamy predyspozycji, aby wiedzieć, jak mało wiemy”. Dan Arieli z kolei porównuje błędy poznawcze i nasze iluzje codziennych wyborów do iluzji optycznych. Dodaje jednak, że te drugie możemy łatwo zmierzyć. W przypadku tych pierwszych – iluzji wyboru – nie mamy tak prostego sposobu weryfikacji.

Ale dlatego właśnie, że jesteśmy ślepi na naszą ślepotę na błędy, powinniśmy przede wszystkim być mniej ufni. I powinniśmy zadawać sobie pytanie – dlaczego wierzę w to, w co wierzę.

Na koniec: dlaczego wciąż się wzruszam na myśl o bezdomnych lamach?

Przyszedł kiedyś do mnie mój student i powiedział, że w sali jest zwierzę. No raczej, mówię, sceniczne, to o mnie, to ja, a on mówi, że nie, że to jest raczej takie zwierzę, co mu bliżej do ptaka. No więc wchodzę do sali i widzę: po prawej studenci zbili się w jedną masę jak gigantyczna bajaderka, po lewej zaś, kulturalnie, na ławce, przysiadł sobie gołąb.

No dobra – mówię nadzwyczaj spokojnie jak na rozmiar tego kryzysu – czy ktoś słyszał kiedyś o gołębiu, który rzucił się na człowieka i wydziobał mu twarz? Poważnie pytam, bo jeśli ktoś słyszał, to ja natychmiast wychodzę, bo ja się panicznie boję śmierci, ale jeśli nikt nie słyszał, to chętnie zostanę bohaterem i zaryzykuję, ale musicie wiedzieć, że ja tak chojraczę tylko dlatego, że to jednak jest gołąb. Co innego, gdyby to była mewa, mewy są chore psychicznie, jak taka na ciebie patrzy, to tam tylko oczy i serce puste, i człowiek od razu wie, że nie miałaby skrupułów. Pamięci wszystkich tych, którzy polegli w boju z mewą, co chciała ich frytkę: [*]

No dobra, nikt nie słyszał, nikt w rodzinie nie ma wujka, który zginął po spotkaniu z sadystycznym ptakiem albo musiał handlować życiem za kawałek precla. Kolega spod okna sprawdził i internet też milczy, oczywiście oprócz tych przezabawnych zdjęć gołębi, co noszą na sobie kromki chleba jak korale, mój Boże, łyzy jak groch ocieram, ja nie wiem, z czego ludzie się kiedyś śmiali, jak nie było Google'a albo gorzej – gołębi.

Gołąb zaś wciąż sobie siedzi, więc ja włączam komputer, Power Pointa, a moi studenci nic, nieruchomo tam w tym kącie jak ta bajaderka, więc ich pytam, czemu długopisu nie wyjmą, karteczki, ja nie mówię, że wszystko mają zanotować, ale chociaż kilka moich brawurowych żartów dla przyszłych pokoleń, ewentualnie naszkicować mogą jakiś tam mój portret. A oni mi mówią, że nie, że oni nie mogą się skupić, jak ten gołąb tak na nich patrzy, no bo co on tak patrzy. I wtedy jeden ochotnik bohater mówi, że on to by może spróbował tego gołębia wyprosić. Znaczy jak? – pytam. „Przepraszam, panie gołębiu, późno już, czy byłby pan tak łaskawy i opuścił nasz lokal?”. A on mówi, że nie, że on by tak gruchał, tak gruchałby i wtedy ten gołąb odleci. A ja mówię, że to świetny pomysł,

niech mi tylko powie, w ramach klasyfikacji poziomów biegłości językowej Rady Europy, to na jakim poziomie on grucha, czy B2, C1, czy wyżej, a on nie odpowiada, to znaczy przynajmniej nie odpowiada mnie, bo gołębiowi jak najbardziej, to znaczy, on mówi: GRUUUUU.

GRRRUUUUUUUUUUU.

Gołąb nic.

GRRRRUUUUUUUUUUUUUU.

Reszta sali też się zainteresowała, obserwujemy globalny rozpad bajaderki, teraz już wszyscy robią GRUUUUU, a ten gołąb najpewniej myśli sobie, że jak wróci do kolegów gołębi i im opowie, co dzisiaj przeżył, to przecież żaden mu w to nie uwierzy. No i ten gołąb sobie siedzi, a moi studenci, wszyscy, jak jeden mąż: GRRUUUUU. Halo, przepraszam, czy jest na sali psycholog społeczny, czy ktoś mógłby zadzwonić do profesora Zimbardo?! I oni tak sobie to gru, a gołąb nic, drzwi otwarte na oścież, jakby chciał wylecieć, i jeszcze na wypadek, gdyby gdzieś zostały mi resztki autorytetu wśród kolegów i koleżanek z wydziału.

Dobra, myślę sobie, gołąb ewidentnie wszystkich ignoruje, to najpewniej jest gołąb anarchista występujący przeciwko mechanizmom wyzysku kapitalistycznej gospodarki, on zmęczony jest codzienną walką o każdy okruszek chleba, oto dowódca gołębiego proletariatu, a gdzie indziej znajdzie schronienie przed wyzyskiem klasy robotniczej jak nie w bezpiecznych murach uniwersytetu. Uwaga – oświadczam – ignorujemy gołębia. Uwaga, zaczynam zajęcia. Puszczam pierwszy slajd, puściłam, gołąb patrzy, patrzy, przemyślał, odleciał.

Człowiek z pierwszej ławki podnosi rękę. On podnosi rękę i ja już wiem, ja nie mam wątpliwości, że wcale nie chce spytać o tematykę zajęć.

Janina – pyta mnie, a bardzo jest zadowolony – czy nie jest ci smutno, że ten gołąb zobaczył twoje slajdy i uciekł?

Nie chcę tego komentować. To naprawdę był bardzo interesujący slajd. I tylko drobne potknięcie na mojej ścieżce pedagogicznego geniuszu, utkanej z pluszowych nitek wiedzy, dobrej zabawy i momentów słodkich niczym małe zajaczki o okrągłych polikach i śmiesznie ruszających się noskach. Na przykład wtedy, kiedy to moi studenci cieszyli się z tego, że coś zrozumieli, a czasem to nawet zainteresowało ich jakieś dodatkowe zagadnienie i pytali: ej, Janina, dlaczego skoro mamy *pie chart* i *donught chart*¹²³, to nigdzie nie ma *cheesecake chart*?

Powiem jednak szczerze, że nie zawsze miałam chwilę pozastanawiać się nad koncepcją wykresu sernikowego. Od pewnego czasu byłam bowiem zajęta zupełnie innymi pytaniami, a mianowicie próbowałam rozwikłać, w jaki sposób doszliśmy do momentu, w którym w kwestii odkryć naukowych, wniosków dotyczących kształtu ziemi, GMO czy różnorodnych rodzajów promieniowania, a także różnorodnych wyników zaawansowanych badań medycznych – również tych dotyczących własnego zdrowia i zdrowia naszych bliskich – wierzymy bardziej inżynierowi górnictwa i agentce ubezpieczeniowej bez jakiegokolwiek wykształcenia medycznego niż nauce.

Przyznaję, niegdyś było we mnie mnóstwo pogardy. Przyjęłam za rzecz oczywistą, że wszyscy mają taką samą wiedzę o badaniach klinicznych i związkach przyczynowo-skutkowych. Uznałam za oczywiste, że wiara w naukę i naukowców powinna być w jakiś sposób automatyczna, samowystarczalna, że nie musimy nikomu tłumaczyć tego, co robimy i dlaczego, niech świat po prostu przyjmie, że (prawie) zawsze mamy rację.

Chwilę zajęło mi przyznanie samej przed sobą, że tak, mamy kryzys wiary w naukę. I tak, my, naukowcy częściowo sami na to zapracowaliśmy. Że musimy zejść z piedestału i zacząć mówić językiem zrozumiałym dla wszystkich niczym mój student mówiący językiem gołębi. Musimy zacząć się komunikować w sposób pozbawiony pogardy, nawet w przypadku najdziwniejszych teorii, niepewności i pytań. **WSZAK HISTORIA NIE ZNA TAKICH PRZYPADKÓW, ŻE KTOŚ SIĘ CZEGOŚ NAUCZYŁ, BO GO UPOKORZYLIŚMY.**

Trochę mi zajęło uzmysłowienie sobie, że pogarda i wrzucanie śmiesznych screenów o dziwnych teoriach antyszczepionkowców lub płaskoziemców niczego nie zmieniają. **NIE MOŻEMY JEDNOCZEŚNIE KOGOŚ OBRAŻAĆ I DZIWIĆ SIĘ, ŻE NIE CHCE NAS SŁUCHAĆ.** Zresztą pomyślcie o tym inaczej: gdy my wrzucamy kolejny śmieszny mem o matce, która pisze, że dziecku z czterdziestostopniową gorączką trzeba przetrzeć twarz zużytymi majtkami, he, he, i na tym kończymy swoje działanie – na wrzuceniu tego screena i rozkosznym rehotaniu, to dokładnie w tej samej chwili jest gdzieś dziecko z czterdziestostopniową gorączką, które nie otrzymuje pomocy.

A przecież w gruncie rzeczy zawsze tłumaczyłam moim studentom, że to nie wstyd, że czegoś nie wiedzą. Że ich najważniejszym nawykiem ma być podejrzliwość w stosunku do wszystkiego, co czytają, i do wszelkich

wniosków, które wyciągają. Że zawsze mają pytać, czy to badanie ma sens, czy też jest fuszerką podobnie jak ten remont, co to w 1995 zrobił nam pan Mirek. Wiecie, ten, który kafelki położył całkiem spoko, z tym że w salonie zamiast w kuchni. I jeszcze na suficie. Mam nadzieję, że wam też się to przytrafi po lekturze tej książki. W sensie – te pytania i ta podejrzliwość do wszelkich doniesień naukowych, nie że remont sygnowany imieniem pana Mirka, jego kreatywnością i umiłowaniem do bimbru sąsiada. To nie wstyd, że czegoś nie wiemy. Problem zaczyna się wtedy, gdy przestajemy pytać.

Choć oczywiście były chwile, kiedy bardzo żałowałam, że ich tak zachęcałam do krytycznego myślenia i zadawania pytań. Na przykład wtedy, gdy przychodzili do mnie zgłębiać wiedzę o bezdomnych lamach biegających po Bydgoszczy.

Studenci na krańcu świata i ja na granicy prawa, a lamy w Bydgoszczy

Zacząło się od tego, że kiedyś zawodowe życie rzuciło mi trochę okruchów suchego chleba pod nogi, kwaśną pomidorową rzeczywistości doprawiło radością i pogodą ducha, co oznacza mniej więcej tyle, że odwołano moje własne zajęcia.

To dlatego, że wszystkim nam kazali opuścić sale wykładowe, bo akurat na uczelni kręcili film wojenny i psuliśmy im kadry, to znaczy, że podobno w czasie wojny ludzie owszem, nosili adidas, ale tylko takie zrobione z rzemyków, a iPhone'ów używali rzadko, bo był słaby zasięg, i tylko niekiedy pozdrawiali mamę do kamery. Moja osobista teoria jest jednak trochę inna i stanowi, że to wszystko przez zasadę kontrastu i że przy nauczycielach, najsmutniejszych ludziach świata, postaci rannych żołnierzy nagle stały się nad wyraz pogodne. Nie, serio, spróbujcie kiedyś komuś po raz sto trzydziesty ósmy bezskutecznie tłumaczyć, jak liczyć prawdopodobieństwo, jednocześnie nogą odpychać studenta, który zasnął i sturlał się pod katedrę, i jeszcze drzeć się do Rudego spod okna, żeby natychmiast przestał udawać pirata z lornetą zrobioną z tablic statystycznych. Jeśli to nie jest wojna, to ja nie wiem, co nią jest.

W każdym razie mnie nie trzeba było do idei odwołanych wykładów dwa razy namawiać, pierwsza skakałam przez rannych żołnierzy i kukłę martwego konia i galopowałam do biura z krótkim postojem na cynamonki w bufecie. No to mam kilka dodatkowych godzin. Co zrobić z tym czasem,

co zrobić z tym czasem? Pomyślałam: a, zostanę częścią zorganizowanej grupy przestępczej.

Dokładniej rzecz ujmując, to nasza przestępcza szajka żyjąca na marginesie prawa odpowiedzialna była za działania o charakterze rewolucyjnym, wręcz paraliżującym społeczną wyobraźnię, co oznacza mniej więcej tyle, że tego dnia postanowiliśmy zmienić żarówkę w biurze. Bo trzeba wam wiedzieć, że u nas na uczelni nie wolno samemu zmieniać żarówki, bo to gwałci wszelkie zasady BHP i może prowadzić do poważnych uszkodzeń ciała, a w niektórych przypadkach również do zagrożenia życia. Procedury wprowadzone w trosce o nasze bezpieczeństwo stanowią, że jeśli się chce w biurze zmienić żarówkę, to należy wezwać miłego pana, który to miły pan ma wszelkie uprawnienia do zmiany żarówki w taki sposób, by zminimalizować liczbę ofiar śmiertelnych i przypadkowych rannych. My jednak postanowiliśmy zmienić żarówkę sami, bo z nas takie krejzole i totalne YOLO, jak to mówią nasi studenci, z tym że YOLO jesteśmy zupełnie od niedawna, to jest od czasu, jak w końcu wygooglowaliśmy, co to znaczy.

Było więc tak, że kolega z ekonomii zmieniał żarówkę, a ja mu pomagałam i to była najbardziej szalona i niezgodna z prawem rzecz, jaką robiłam w tym kraju, przynajmniej od czasu, jak kiedyś w Tesco próbowałam kupić nielegalną ilość paracetamolu. No i był jeszcze kolega z teologii, który pilnował drzwi, czy przypadkiem ktoś nie nadchodzi, bo zgodnie uznaliśmy, że jemu najbardziej przyda się ćwiczenie cierpliwości w wypatrywaniu tego, co być może nigdy nie nadejdzie. I cała operacja zmierzała już ku końcowi, i już cali byliśmy podjarani naszą nową tożsamością ludzi poza prawem, już przygotowywaliśmy się do zakupu papugi na ramię i opaski na oczy, my, piraci zasad społecznych, kowboje regulaminów, miglance paragrafów, a wtedy to kolega spod drzwi wszczął alarm i *THEY'RE COMING! THEY'RE COMING!* – krzyczał, a krzyczał tak dramatycznie, że aż przez chwilę nie byłam pewna, czy widzi pana dziekana, czy też coś znacznie gorszego typu sąd ostateczny albo ciasteczka bezglutenowe bez cukru, ale okazało się, że sprawa jest jeszcze poważniejsza, bo oto nadchodzili studenci.

W takiej sytuacji to zawsze strach, bo wielu studentów mamy wspólnych i to nigdy nie wiadomo, do kogo przychodzą, więc z nas wszystkich tylko kolega z ekonomii powiedział, że on się w ogóle nie boi, bo na pewno nie przyszli do niego, no bo on swoim studentom powiedział, że przebywa

w Ameryce Południowej, a ja go pytam, po co aż w Ameryce Południowej, czy nie mógł po prostu wybrać się do Cork, a on powiedział, że Cork jest zbyt blisko, w Cork mogliby go znaleźć, więc wybrał Amerykę Południową i może być też tak, że wspomniał coś o dżungli. OK, bez paniki. Może wcale nie szukają nikogo z nas, może po prostu się zgubili, szukają biblioteki może? Zaśmiałam się z własnego dowcipu, ale krótko, bo oto nasz odźwierny już szedł przekazać informację, kogo właściwie szukają studenci. Życie, ty takie przewidywalne... jak nie wiesz, kogo wychłostać, to wiadomo, że wybierzesz Polaka.

No to idę. Idę i pytam, czego ode mnie potrzebują, a oni mówią, że chcieli porozmawiać o metodach nieparametrycznych, co mnie szalenie wzrusza, ale tylko przez chwilę, bo zaraz potem trzy lata pedagogicznego doświadczenia dają mi w twarz, odzyskuję trzeźwość i mówię, żeby nie udawali, przecież ich to w ogóle nie interesuje, a ten pierwszy poddaje się dość szybko i mówi, że OK, no dobra, w ogóle go to nie interesuje, a ten drugi jeszcze trochę walczy i mówi, że jednak trochę go to interesuje, ale potem zastanawia się chwilę i mówi, że jednak nie, wcale go to nie interesuje, i że tak właściwie to są tu tylko z jednego powodu, mianowicie takiego, że wybierają się wkrótce na zagraniczny wolontariat. Wiecie, o co chodzi – o te programy, że ludzie jeżdżą do tych mniej fortunnych zakątków świata, by nieść kaganek oświaty, uczyć demokracji, podzielić się odrobiną technologicznych osiągnięć. Więc ja mówię, że to świetny pomysł, i pytam, w jakim to kraju Trzeciego Świata będą łątać dziury cywilizacyjne, czy będzie to Tajlandia? Mozambik? Indie? A oni mówią, że prawie, bo Polska.

Więc pytam, co dokładnie będą w tej Polsce porabiać, a oni mówią, że za dnia uczyć dzieci angielskiego, a w nocy wyrwać na to lachony, wiadomo, wszyscy zostaliśmy nauczycielami z tych samych powodów. I pytają, czy mam jakieś wskazówki na ich nową polską drogę życia, a ja im mówię, żeby przede wszystkim przed tym wyjazdem zbadali sobie wątrobę, i oni zapisują: *CHECK. MY. LIVER.*, oczywiście, że zapisują, a jak im sto razy tłumaczę, co to jest istotność statystyczna, to zapominają zapisać. I jeszcze mówią, że dwa tygodnie temu na zajęciach powiedziałam – i tu cytują – że *I'm from a wild country* i oni chcieliby wiedzieć, czy ja miałam na myśli dzikość dosłowną, w sensie, czy grozi im śmierć w paszczy niedźwiedzia, czy też to była taka metafora, a ja mówię, że nigdy nic takiego nie powiedziałam, a oni mi pokazują zeszyt i okazuje się, że 9 kwietnia

rzeczywiście coś takiego powiedziałam. Generalnie doskonale wiedzieć, że oni takie rzeczy zapisują i że już nigdy nie zostanę prezydentem.

No więc pytają mnie o ten nasz dziki kraj, to jest – na ile jest dziki i jak bardzo muszą się przygotować, czy wystarczy im zwykła kurtka, czy muszą kupić taką na wyprawę polarne, a ja ich pytam, skąd im się wzięła ta polarność, a oni mówią, że z nazwy, no bo POLA-ND, to trochę jak POLA-R, i jeszcze pytają, czy dużo jest w tej Polsce *polar bears*, a ja mówię, że trochę, ale to raczej na wsi, już coraz rzadziej się je widuje w miastach, i oni to też zapisują.

Szczerze mówiąc, to nie liczyłam już za bardzo na to, że złapią się na te niedźwiedzie, zwłaszcza że dwa lata temu mieliśmy podobną dyskusję, kiedy to rozmawiałam ze studentami o irlandzkich problemach społecznych i rzuciłam brawurowy żart o tym, że obecnie w Polsce największym problemem społecznym są bezpańskie lamy wałęsające się po ulicach, i to zupełnie ich przerosło, w ogóle nie mieściło im się w głowie, no bo dlaczego rząd na to pozwala, na te stada agresywnych lam biegających po ulicach, dlaczego, spytali, nikt w Polsce nie wybuduje więcej schronisk dla bezdomnych lam?! Niemniej w tamtym momencie okazało się, że ta dziwna kombinacja zwierząt w ogóle nie wzbudziła niczyich podejrzeń, wiadomo, dzień jak co dzień w Polsce – niedźwiedzie polarne bawią się z lamami, a małpy poganiają je od tyłu.



Mój Boże, weźcie sobie wyobraźcie ten tragiczny krajobraz miejski, te stada bezdomnych lam wałęsające się po Bydgoszczy! Ja do dziś lubię sobie czasem ten obraz przypominać, bo dla was to tylko historia o dzikich zwierzętach pozbawionych schronienia i znizek w komunikacji miejskiej największych polskich miast, a dla mnie wspomnienia ciepłe niczym ogrzewanie w PKP w samym środku lata. Bo gdy w końcu wyjaśniliśmy sobie wszystkie zagadnienia związane z polską fauną i florą, to moi studenci powiedzieli coś, co było dla mnie wydarzeniem na miarę pierwszego słowa dziecka (moment dla mnie abstrakcyjny, ale totalnie potrafię go sobie wyobrazić, wiecie, tę chwilę, gdy ten nieporadny niemowlak po raz pierwszy łapie rodzica za palec i mówi: kowariancja).

Prawdą jest, że w tamtym momencie moi studenci nie powiedzieli: kowariancja. Powiedzieli natomiast, że jest jedna, najważniejsza sprawa, która ich od kilku dni gnębi, to znaczy chcieliby wiedzieć, czy to prawda, że kobiety w Polsce są taaaaaaakie ładne, ale to taaaaaaakie ładne? A ja im na to powiedziałam, że w sumie to nie wiem, to znaczy na przykład mój mąż twierdzi, że ja jestem bardzo ładna, zwłaszcza gdy chce, bym mu skoczyła do Żabki po browara, ale musimy być ostrożni z generalizacją wyników tego badania, a jeden z nich mówi, że wie, wie, przecież pamięta z moich zajęć, że w celu generalizacji musielibyśmy pobrać dostatecznie liczną próbę w sposób losowy.

OK, może ja powtórzę, żebyście nie przegapili:

POWIEDZIAŁ, ŻE PAMIĘTA COŚ Z MOICH ZAJĘĆ.

I wtedy zorientował się, co powiedział, i zamilkł.

I ja też milczałam.

I razem milczeliśmy.

Po policzku spływała mi łza.

W tle słychać było bicie dzwonów.

Aż w końcu on poczuł się w obowiązku wytłumaczyć: *I have no idea how I learned that but I better drink it away before I remember it forever.*

I poszedł. A jak tylko zniknął za rogiem, to pobiegłam do biura ile sił w nogach i powiedziałam mojemu koledze z biurka obok, że właśnie przed chwilą okazało się, że jeden ze studentów się czegoś na moich zajęciach nauczył, i koleżance obok też powiedziałam, i koledze spod okna też. Nikt mi nie uwierzył. Wiadomo, tych Irlandczyków to nie tak łatwo nabrać. A nauczony student to nie niedźwiedź polarny ujeżdżany przez tapira na warszawskiej Starówce – występuje raczej rzadko.

Któż po latach pamięta o małpach poganiających tapiry w Kaliszu?

Nie mam pojęcia, czy kilka lat po otrzymaniu dyplomu moi studenci wciąż pamiętają, o co chodzi w generalizowaniu wyników badań i czy na ostatniej wycieczce do Wieliczki wciąż chowali się za winklem ze strachu przed nieoswojonymi ssakami parzystokopytnymi z rodziny wielbłądowatych. Niemniej totalnie wiem, jak skończyła się ich wyprawa do polskiego świata, wszak w kolejnym roku akademickim miałam okazję spotkać się z naszymi gwiazdami wolontariatu. Dokładnie tymi samymi ludźmi, którzy pojechali łątać dziury cywilizacyjne w krajach Trzeciego Świata, to jest mieszkali z polskimi rodzinami w podkrakowskiej wsi, gdzie uczyli angielskiego uzbrojeni w harpuny, ale te harpuny to nie przeciwko dzieciom, tylko przeciwko niedźwiedzim polarnym i dzikim lamom. I całe lato polowali na te lamy, a przy okazji ślali mi zdjęcia siebie i pierogów, siebie w chustach góralskich, siebie szydełkujących tradycyjne serwety na stół, siebie na warsztatach ubijania masła, generalnie siebie w trakcie wykonywania typowych polskich czynności dnia codziennego. A jak już moi studenci wrócili ze tej wyprawy życia, to postanowili obdarować mnie wszystkimi nowo poznanymi polskimi frazami, więc pierwsze pięć minut zajęć spędzili na powtarzaniu mi: *Czeszcz, pyenkna staruvka, gdzie jest tramvay?* I czasem jeszcze: *Mmmmmmm smacne yayko.*

Na co ja powiedziałam, że to wspaniale, że wszyscy się zgadzamy co do tego, że yayko jest smacne, a staruvka pyekna, ale proponuję skupić się na temacie dzisiejszych zajęć i zapytałam ich, czy pamiętają z zeszłego roku, czym różni się regresja logistyczna od probitowej. A wtedy człowiek spod okna wybuchnął śmiechem i powiedział, że to doskonały dowcip, jeden z lepszych moich żartów, a ja powiedziałam, że przecież nie powiedziałam żadnego dowcipu, a on powiedział, że w takim razie przeprasza, on myślał, że moje pytanie o to, czy pamiętają coś z zeszłego roku, to był żart. I wszyscy się śmieją, oprócz tego studenta z drugiego rzędu, naszego podróżnika po polskiej dżungli, który jest zajęty tłumaczeniem czegoś ważnego swojemu sąsiadowi: *Seriously, Rory* – tłumaczy mu – *smacne yayko.*

– *ENOUGH WITH THE YAYKO* – denerwuję się.

– *But... yayko...* – posmutniał nagle nasz polski wędrownik.

– *What is this yayko, anyway?* – zainteresował się człowiek spod okna.

– *'Yayko' means 'egg'* – tłumaczy mu nasz specjalista od kultury polskiej i jeszcze dodaje: – *Polish people eat a lot of yaykos.*

– *They have better yaykos than we have* – mówi jego towarzysz podróży – *because they like kury and treat them well.*

– *Kuuuuu. Ry.* – powtarza ten z pierwszego rzędu. – *That's a funny sound*

– *'Kury' is 'hen'* – tłumaczy mu ten drugi.

– *'Kury' is plural, it's a lot of hens* – poprawiam go.

– *A lot of kury means a lot of yaykos* – przeprowadził skomplikowaną operację logiczną ten pierwszy.

– *You should be really proud of yourself* – szepcze ze wzruszeniem drugi miłośnik polskich jajek.

– *You're a great nation with great yaykos.*

– *And great language* – dorzuca ten drugi i dodaje: – *Gdzie jest tramway?*

– *Kuuuuu. Ry.* – mówi człowiek spod okna.

– *Kury make yaykos.* – Człowiek z pierwszego rzędu podsumował wszystko, czego się nauczył na tamtejszych zajęciach.

– *Valensa.* – Koleżanka z pierwszego rzędu też przypomniała sobie polskie słowo.

– *Kury, yaykos, Valensa, tramway* – streścił nasz kraj miły człowiek spod okna.

– *I like this class a lot* – oceniła nasza świeżo przybyła francuska koleżanka z Erasmusa, dla której okazały się to pierwsze zajęcia na irlandzkiej uczelni. – *I think this will be my favourite one.*

No, dobrze, że ktoś. Ja byłam trochę mniej zadowolona. Czyli jednak – odróżniania zmiennej zależnej od niezależnej to przez trzy lata nie potrafią się nauczyć, ale słownictwo z zakresu polskiego rolnictwa opanowują w piętnaście minut. No to są jakieś yaykos.

Niemniej to wcale nie przez te lamy, yaykos i trudności moich studentów z trzymaniem się tematu odeszłam z uczelni w 2017 roku. W ogóle nie przez studentów. W gruncie rzeczy lubię myśleć, że żyło mi się z nimi dobrze. Owszem, mieliśmy parę trudnych chwil, na przykład wtedy, kiedy to opowiedziałam brawurowy dowcip o Spice Girls, a oni nie tylko się nie zaśmiali, ale co gorsza spytali mnie, kto to jest Spice Girls, co sprawiło, że kolejne cztery tygodnie spędziłam w łóżku, opłakując swoją młodość i dzwoniąc do Telezakupów Mango, by dowiedzieć się, czy ten krem z wyciągiem ze śluzu ślimaka to naprawdę redukuje wszystkie zmarszczki, nawet te, które jeszcze nie powstały. Albo wtedy, kiedy tłumaczyli mi, że według danych ankietowych 350% studentów ma bardzo dobre zdrowie i choć te liczby mogą cieszyć, to należy je traktować z pewną ostrożnością,

albowiem wśród respondentów mamy też 450 i pół osoby, które cierpią na depresję, co mnie w sumie nie dziwi, bo ja też bym miała depresję, gdyby mi brakowało, na przykład, głowy. Tak, mieliśmy ze studentami swoje trudne momenty, ale przez większość czasu żyliśmy w krainie wiecznej szczęśliwości, tj. biegaliśmy wspólnie po akademickiej łące wiedzy, trzymając się za ręce, nucąc *2 become 1* z repertuaru Spice Girls (czegoś o świecie ich jednak musiałam nauczyć) i tropiąc wszelkie fałszywe wnioski, zmanipulowane wykresy, błędne badania naukowe, no i niektórzy z nas jeszcze lamy, wiadomo. Bardzo mnie wzruszają te wspomnienia.

Albowiem jest w tym tropieniu naukowych zagadek coś ekstremalnie satysfakcjonującego i fajnego, tylko odrobinę mniej pluszowego niż ten moment, gdy mój szef przyniósł mi dwanaście muffinów z borówkami pozostałych po konferencji, i trochę bardziej niż ten, gdy w połowie konsumpcji powiedział mi, że one są z zeszłego tygodnia. Na koniec tej książki nie pozostaje mi więc nic innego jak życzyć wam i sobie właśnie takich intelektualnych rozkoszy, mnóstwa rozwiązań naukowych zagadek i poszukiwań prawdy owocnych jak te muffinki (he, he, he).

Czy wiecie, że istnieje matematyczny wzór na dobre piwo?

Drydak, N. (2015), *The effect of sexual activity on wages*, „International Journal of Manpower” 36 (2), 192–215.

Skopek, J., Schulz, F., & Blossfeld, H.P. (2010), *Who contacts whom? Educational homophily in online mate selection*, „European Sociological Review” 27 (2), 180–195.

<https://www.fronda.pl/a/kuchnia-antydepresyjna-omlet-z-warzywami-pieczarkami-i-szynka,81970.html>

ILU osób potrzeba, by napisać jedną książkę?

Westgate, E.C., & Wilson, T.D. (2018), *Boring thoughts and bored minds: The MAC model of boredom and cognitive engagement*, „Psychological Review” 125 (5), 689.

Statystyka – królowa nauk, czy popychadło medialnego świata?

Westgate, E. C., & Wilson, T. D. (2018). *Boring thoughts and bored minds: The MAC model of boredom and cognitive engagement*. „Psychological review”, 125(5), 689.

Czy naukowcy są jak żony i nigdy się nie mylą?

Anderson, P.C. (1976), *The Patchwork Mouse*, „JAMA” 236 (2), 205–205.

Blaustien, M. (2015), *Academic paper on plagiarism pulled because of plagiarism*, dostępny on-line: <https://nypost.com/2015/04/02/academic-paper-on-plagiarism-pulled-because-of-plagiarism/> (dostęp 20.01.2020).

Boddy, J. (2016), *A cat co-authored an influential physics paper*, dostępny on-line: <https://www.sciencemag.org/news/2016/08/cat-co-authored-influential-physics-paper> (dostęp 20.01.2020).

Campos-Arceiz, A., Primack, R.B., & Koh, L.P. (2015), *Reviewer recommendations and editors' decisions for a conservation journal: Is it just*

- a crapshoot? And do Chinese authors get a fair shot?*, „Biological Conservation” 186, 22–27.
- Carter, K. i Carter, B. (2005). *Childbed Fever: A Scientific Biography of Ignaz Semmelweis*. New York: Routledge.
- Dubner, S. J., & Levitt, S. D. (2010). *Superfreakonomics: Global Cooling, Patriotic Prostitutes and Why Suicide Bombers Should Buy Life Insurance*, Penguin UK.
- Edmonds, D., Eidinow, J., (2002), *Pogrzebacz Wittgensteina. Opowieść o dziesięciominutowym sporze między dwoma wielkimi filozofami*, przeł. Lech Niedzielski, Warszawa: Muza.
- Farrington, C.P., Miller, E., & Taylor, B. (2001), *MMR and autism: further evidence against a causal association*, „Vaccine” 19 (27), 3632–3635.
- Gadad, B.S., Li, W., Yazdani, U., Grady, S., Johnson, T., Hammond, J., Ferrier, C. (2015), *Administration of thimerosal-containing vaccines to infant rhesus macaques does not result in autism-like behavior or neuropathology*, „Proceedings of the National Academy of Sciences” 112 (40), 12 498–12 503.
- Helmer, M., Schottdorf, M., Neef, A., & Battaglia, D. (2017), *Gender bias in scholarly peer review*, „Elife” 6, e21718.
- Hviid, A., Hansen, J.V., Frisch, M., & Melbye, M. (2019), *Measles, Mumps, Rubella Vaccination and Autism: A Nationwide Cohort Study*. *Annals of internal medicine*.
- Hewitson, J. (2018), *Autism: how to raise a happy autistic child*, s. loc. 232, section: Andrew Wakefield.
- Mrozek-Budzyn, D., Kieltyka, A., & Majewska, R. (2010), *Lack of association between measles-mumps-rubella vaccination and autism in children: a case-control study*, „The Pediatric infectious disease journal” 29 (5), 397–400.
- Obenchain, T. G. (2016). *Genius Belabored: Childbed Fever and the Tragic Life of Ignaz Semmelweis*. Tuscaloosa: University of Alabama Press.
- OECD (2019), *Child vaccination rates*, dostępny on-line: <https://data.oecd.org/healthcare/child-vaccination-rates.htm> (dostęp 20.01.2020).
- Spektor, B. (2018), *The author of this physics paper is 7 years old (and also a cat)*, dostępny on-line: <https://www.livescience.com/61698-cat-authors-physics-paper.html> (dostęp 20.01.2020).
- Wakefield, A.J., Murch, S.H., Anthony, A., Linnell, J., Casson, D.M., Malik, M., Valentine, A. (1998), *Retracted: Ileal-lymphoid-nodular hyperplasia*,

non-specific colitis, and pervasive developmental disorder in children.

Westling, T. (2011), *Male organ and economic growth: does size matter?*,
dostępny on-line:
<https://helda.helsinki.fi/bitstream/handle/10138/27239/maleorga.pdf>
(dostęp 20.01.2020).

WHO (2014) *Global Immunization Data*, dostępny on-line:
https://www.who.int/immunization/monitoring_surveillance/global_immunization_data.pdf (dostęp 20.01.2020).

<https://www.tylervigen.com/spurious-correlations>

<https://ecdc.europa.eu/en/home>

Dlaczego koń i szop pracz nie dostały doktoratu?

Abadie, A., & Gardeazabal, J. (2003), *The economic costs of conflict: A case study of the Basque Country*, „American economic review” 93 (1), 113–132.

Bayou, M.E., & De Korvin, A. (2008), *Measuring the leanness of manufacturing systems—a case study of Ford Motor Company and General Motors*, „Journal of Engineering and Technology Management” 25 (4), 287–304.

Cairns, E., & Darby, J. (1998), *The conflict in Northern Ireland: Causes, consequences, and controls*, „American psychologist” 53 (7), 754.

Fielding, D. (2003), *Investment, employment, and political conflict in Northern Ireland*, „Oxford Economic Papers” 55 (3), 512–535.

Government of Odisha (2010), *Census in different countries. Orrisa review (census special)*, dostępny on-line:
<http://magazines.odisha.gov.in/Orissareview/2010/December/engpdf/26-36.pdf> (dostęp 20.01.2020).

Hoelzel, A.R. (1999), *Impact of population bottlenecks on genetic variation and the importance of life-history; a case study of the northern elephant seal*, „Biological Journal of the Linnean Society” 68 (1–2), 23–39.

Mann, K.H. (1965), *Energy transformations by a population of fish in the River Thames*, „The Journal of Animal Ecology”, 253–275.

Strong, P.N. (1999), *The economic consequences of ethno-national conflict in Cyprus: the development of two siege economies after 1963 and 1974 (Doctoral dissertation)*, The London School of Economics and Political Science).

Warren, J.M., & Warren, H.B. (1962), *Reversal Learning by Horse and Raccoon*, „The Journal of Genetic Psychology” 100 (2), 215–220.

Skąd wiemy, że owca rozpoznałaby na ulicy prezydenta?

Ghirlanda, S., Jansson, L., & Enquist, M. (2002), *Chickens prefer beautiful humans*, „Human Nature” 13 (3), 383–389.

Kendrick, K.M., Atkins, K., Hinton, M.R., Broad, K.D., Fabre-Nys, C., & Keverne, B. (1995), *Facial and vocal discrimination in sheep*, „Animal Behaviour” 49 (6), 1665–1676.

Kendrick, K.M., Atkins, K., Hinton, M.R., Heavens, P., & Keverne, B. (1996), *Are faces special for sheep? Evidence from facial and object discrimination learning tests showing effects of inversion and social familiarity*, „Behavioural Processes” 38 (1), 19–35.

Knolle, F., Goncalves, R.P., & Morton, A.J. (2017), *Sheep recognize familiar and unfamiliar human faces from two-dimensional images*, „Royal Society open science” 4 (11), 171–228.

Peirce, J.W., Leigh, A.E., & Kendrick, K.M. (2000), *Configurational coding, familiarity and the right hemisphere advantage for face recognition in sheep*, „Neuropsychologia” 38 (4), 475–483.

Peirce, J.W., Leigh, A.E., & Kendrick, K.M. (2001), *Human face recognition in sheep: lack of configurational coding and right hemisphere advantage*, „Behavioural Processes” 55 (1), 13–26.

Smith, M.L. (2014), *Honey bee sting pain index by body location*, „PeerJ” 2, e338.

Czy kiedy wychodzimy z psem na spacer, to średnio mamy po trzy nogi?

Ghirlanda, S., Jansson, L., & Enquist, M. (2002), *Chickens prefer beautiful humans*, „Human Nature” 13 (3), 383–389.

Jeffers, J.N.R. (1998), *The statistical basis of sampling strategies for rivers: an example using River Habitat Survey*, „Aquatic conservation: marine and freshwater ecosystems” 8 (4), 447–454.

Kokko, A., & Tingvall, P.G. (2012), *The Eurovision song contest, preferences and European trade*, „Ratio Working Paper” 183.

Newberry, R.C., Hunt, J.R., & Gardiner, E.E. (1985), *Behaviour of roaster chickens towards an automatic weighing perch*, „British poultry science” 26

(2), 229–237.

Padley, F.B., & Timms, R.E. (1980), *The determination of cocoa butter equivalents in chocolate*, „Journal of the American Oil Chemists' Society” 57 (9), 286–293.

Czy 117% Polaków może się mylić?

Wykres kołowy ze strony 122: <https://starsgalaxiescosmologylab.wordpress.com/> (dostęp 20.01.2020).

Wykres ze strony 128 IPUMS USA 2017; US Bureau of Labor Statistics; McKinsey Global Institute Analysis.

Wykres ze strony 131: pojawił się 15 października 2018 roku na oficjalnym koncju NFZ na Twitterze (@NFZ_Centrala).

Wykres ze strony 132: pojawił się 8 kwietnia 2019 roku na oficjalnym koncju Kancelarii Premiera na Twitterze (@PremierRP).

Wykres ze strony 134: *Betlej: Czy wrzesień w biurach podróży przegoni lipiec?*, dostępny on-line: <https://www.rp.pl/Biura-podrozy/309189997-Betlej-Czy-wrzesien-w-biurach-podrozy-przegoni-lipiec.html> (dostęp 20.01.2020).

Wykres ze strony 135: <https://starpets.us/blue-jays-vs-yankees.html> (dostęp 20.01.2020).

Wykres ze strony 136: http://jimcgreevy.com/gvdc/Natl_Debt_Chart.html (dostęp 20.01.2020).

Wykres ze strony 137: Florida Department of Law Enforcement (opublikowano: 16.02.2014). Autorka: C. Chan.

Cleveland, W.S., & McGill, R. (1985), *Graphical perception and graphical methods for analyzing scientific data*, „Science” 229 (4716), 828–833.

Hollands, J.G., & Spence, I. (1998), *Judging proportion with graphs: The summation model*, „Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition” 12 (2), 173–190.

Simkin, D., & Hastie, R. (1987), *An information-processing analysis of graph perception*, „Journal of the American Statistical Association” 82 (398), 454–465.

Spence, I., & Lewandowsky, S. (1991), *Displaying proportions and percentages*, „Applied Cognitive Psychology” 5 (1), 61–77.

Tufte, E.R. (2001), *The visual display of quantitative information (Vol. 2)*. Cheshire, CT: Graphics press.

<https://www.gapminder.org>
<https://informationisbeautiful.net>

Czy da się zmierzyć szczęście linijką?

Błuszkowski, J. (2006), *Przemiany stereotypów narodowych w procesie transformacji systemowej w Polsce*, „Studia Polityczne” 11, 93–106.

Kendrick, K.M., da Costa, A.P., Leigh, A.E., Hinton, M.R., & Peirce, J.W. (2001), *Sheep don't forget a face*, „Nature” 414 (6860), 165.

Samady, S. (2013), *Changing Profile of Education in Afghanistan*, dostępny on-line: <http://eslus.com/articles/Ed-Afghanistan.pdf> (dostęp 20.01.2020).

Shayan, Z. (2015), *Gender Inequality in Education in Afghanistan: Access and Barriers*, „Open Journal of Philosophy” 5, 277–284, dostępny on-line: <http://dx.doi.org/10.4236/ojpp.2015.55035> (dostęp 20.01.2020).

The campus squirell listings, dostępny on-line: <http://www.gottshall.com/squirrels/campsq.htm> (dostęp 20.01.2020).

Trzebińska, E. (2008), *Psychologia pozytywna*. Warszawa: Wydawnictwa Akademickie i Profesjonalne.

Whoops: Data blunder sees TCD fall out of global rankings, dostępny on-line: <https://www.irishtimes.com/news/education/whoops-data-blunder-sees-tcd-fall-out-of-global-rankings-1.2800587> (dostęp 20.01.2020).

<http://happyplanetindex.org/>

<http://perspektywy.pl/RSW2019/>

<http://www.oecdbetterlifeindex.org/>

http://www.polityka.pl/nauka/_resource/multimedia/20034494

<http://www.worldvaluessurvey.org/wvs.jsp>

<https://ourworldindata.org/happiness-and-life-satisfaction>

<https://worldhappiness.report/>

Czy po seksie będziemy opanierowani w banknotach niczym krokiet
w bułce?

Ackerman, Joshua M., Nocera, Christopher C., & Bargh, John A. (2010), *Incidental haptic sensations influence social judgments and decisions*, „Science” 328.5986, 1712–1715.

- Clariana, R., & Wallace, P. (2002), *Paper-based versus computer-based assessment: key factors associated with the test mode effect*, „British Journal of Educational Technology” 33 (5), 593–602.
- Davies, J.B., & Baker, R. (1987), *The impact of self-presentation and interviewer bias effects on self-reported heroin use*, „British Journal of Addiction” 82 (8), 907–912.
- Egleston, B.L., Miller, S.M., & Meropol, N.J. (2011), *The impact of misclassification due to survey response fatigue on estimation and identifiability of treatment effects*, „Statistics in Medicine” 30 (30), 3560–3572.
- Groves, R.M. (2004), *Survey errors and survey costs vol. 536*, John Wiley & Sons.
- Hart, T.C., Rennison, C.M., & Gibson, C. (2005), *Revisiting respondent fatigue bias in the National Crime Victimization Survey*, „Journal of Quantitative Criminology” 21 (3), 345–363.
- Krumpal, I. (2013), *Determinants of social desirability bias in sensitive surveys: a literature review*, „Quality & Quantity” 47 (4), 2025–2047.
- Marta-Pedroso, C., Freitas, H., & Domingos, T. (2007), *Testing for the survey mode effect on contingent valuation data quality: A case study of web based versus in-person interviews*, „Ecological economics” 62 (3–4), 388–398.
- Mcbee, G.W., & Justice, B. (1977), *The effect of interviewer bias on mental illness questionnaire responses*, „The Journal of psychology” 95 (1), 67–75.
- Rodriguez, S. (2014), *1 in 10 in a survey think HTML is an STD*, „Los Angeles Times” 04.03.2014, dostępny on-line: <https://www.latimes.com/business/technology/la-fi-tn-1-10-americans-html-std-study-finds-20140304-story.html#axzz2v1X0Ct00> (dostęp 20.01.2020).
- Williams, Lawrence E., & John A. Bargh, *Experiencing physical warmth promotes interpersonal warmth*, „Science” 322.5901 (2008), 606–607.

Na imieniny lepiej kupić czekoladki czy kokainę?

- Bruinsma, K., & Taren, D.L. (1999), *Chocolate: food or drug?*, „Journal of the American Dietetic Association” 99 (10), 1249–1256.
- Couper, M.P. (1997), *Survey introductions and data quality*, „Public Opinion Quarterly” 61 (2), 317–338.

- Couper, M.P. (2000), *Web surveys: A review of issues and approaches*, „Public Opinion Quarterly” 64 (4), 464–494.
- Couper, M.P., & Nicholls II, W.L. (1998), *The history and development of computer assisted survey information collection methods*, [in:] M.P. Couper, R.P. Baker, J. Bethlehem, C.Z.F. Clark, J. Martin, W.L. Nicholls II, & J.M. O’Reilly (eds.), *Computer assisted survey information collection* (1–21), New York, NY: John Wiley & Sons.
- Couper, M.P., Traugott, M., & Lamias, M. (2000), *Experiments on the design of Web surveys. Paper presented at the Fifth International Conference on Social Science Methodology, Cologne, Germany, October 3–6, 2000.*
- Couper, M.P., Traugott, M., & Lamias, M. (2001), *Web survey design and administration*, „Public Opinion Quarterly” 65 (2), 230–253.
- Dillman, D.A., Tortora, R.D., & Bowker, D. (1998), *Influence of plain vs. fancy design on response rates for Web surveys. The 1998 Proceedings of Section on Survey Research Methods*, Dallas, Texas: American Statistical Association, [dostępny on-line: http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf](http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf) (dostęp 20.01.2020).
- Dillman, D.A. (2000), *Tailored design of mail and other self-administered surveys*, New York, NY: Wiley-Interscience.
- Dillman, D.A., & Bowker, D.K. (2001), *The Web questionnaire challenge to survey methodologists*, [in:] U.-D. Reips, & M. Bosnjak (eds.), *Dimensions of Internet science* (159–178), Lengerich: Pabst Science Publishers.
- Hu, N., Pavlou, P., & Zhang, J. (2008), *Overcoming Self-Selection Biases in Online Product Reviews*.
- Infante-Rivard, C., & Jacques, L., *Empirical Study of Parental Recall Bias*, „American Journal of Epidemiology” vol. 152, issue 5, 01.09.2000, 480–486.
- Krotki, K.P. (2001), *Web-based surveys for a nationally representative sample. Paper presented at Joint Statistical Meeting 2001, Atlanta, Georgia, 5–9 August 2001.*
- Lenoir, M., Serre, F., Cantin, L., & Ahmed, S.H. (2007), *Intense sweetness surpasses cocaine reward*, „PloS one” 2 (8), e698, [dostępny on-line: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000698](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000698).
- Li, X., & Hitt, L.M. (2008), *Self-selection and information role of online product reviews*, „Information Systems Research” 19 (4), 456–474.

Czy wszyscy jesteśmy rasistami?

- Albert, R., Escot, L., & Fernández-Cornejo, J.A. (2011), *A field experiment to study sex and age discrimination in the Madrid labour market*, „The International Journal of Human Resource Management” 22 (02), 351–375.
- Baert, S., & Verhofstadt, E. (2015), *Labour market discrimination against former juvenile delinquents: Evidence from a field experiment*, „Applied Economics” 47 (11), 1061–1072.
- Bertrand, M., & Mullainathan, S. (2004), *Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination*, „American economic review” 94 (4), 991–1013.
- Bloom, S.G. (2005), *Blue-eyes, brown-eyes: The experiment that shocked the nation and turned a town against its most famous daughter*, „Smithsonian Magazine”, dostępny on-line: <https://blogdomensalao.files.wordpress.com/2013/02/blue-eyes-brown-eyes-bysgbloom.pdf> (dostęp 20.01.2020).
- Bloom, S.G., & Kennedy, L. (2005), *Lesson of a lifetime. Teacher Jane Elliott’s unorthodox exercise to instruct her third graders in the consequences of racism still divides an Iowa town-and even her former students-nearly 40 years later*, „Smithsonian Magazine” 36 (6), 82.
- Blum, B. (2018), *The lifespan of a lie*, „Medium”, dostępny on-line: <https://gen.medium.com/the-lifespan-of-a-lie-d869212b1f62> (dostęp 20.01.2020).
- Booth, A., & Leigh, A. (2010), *Do employers discriminate by gender? A field experiment in female-dominated occupations*, „Economics Letters” 107 (2), 236–238.
- Doliński, D., Grzyb, T., Folwarczny, M., Grzybała, P., Krzyszycha, K., Martynowska, K., & Trojanowski, J. (2017). *Would you deliver an electric shock in 2015? Obedience in the experimental paradigm developed by Stanley Milgram in the 50 years following the original studies*. „Social psychological and personality science”, 8(8), 927-933.
- Drydakis, N. (2009), *Sexual orientation discrimination in the labour market*, „Labour Economics” 16 (4), 364–372.
- Drydakis, N., MacDonald, P., Chiotis, V., & Somers, L. (2018), *Age discrimination in the UK labour market. Does race moderate ageism? An experimental investigation*, „Applied economics letters” 25 (1), 1–4.

- Elliott, J., *A class divided*, film z 1985 roku, będący częścią cyklu Frontline nadawanego przez stację PBS.
- Elliott, J., *The eye of the storm*, krótki film dokumentalny z 1970 roku, wyświetlony po raz pierwszy przez stację ABC.
- González, M.J., Cortina, C., & Rodríguez, J. (2019), *The Role of Gender Stereotypes in Hiring: A Field Experiment*, „European Sociological Review” 35 (2), 187–204.
- Kaas, L., & Manger, C. (2012), *Ethnic discrimination in Germany’s labour market: a field experiment*, „German economic review” 13 (1), 1–20.
- Kang, S.K., DeCelles, K.A., Tilcsik, A., & Jun, S. (2016), *Whitened resumes: Race and self-presentation in the labor market*, „Administrative Science Quarterly” 61 (3), 469–502.
- Le Texier, T. (2019), *Debunking the Stanford Prison Experiment*, „American Psychologist” 74 (7), 823–839.
- McGinnity, F., & Lunn, P.D. (2011), *Measuring discrimination facing ethnic minority job applicants: an Irish experiment*, „Work, employment and society” 25 (4), 693–708.
- Pager, D., Bonikowski, B., & Western, B. (2009), *Discrimination in a low-wage labor market: A field experiment*, „American sociological review” 74 (5), 777–799.
- Petit, P. (2007), *The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector*, „Labour Economics” 14 (3), 371–391.
- Riach, P.A., & Rich, J. (2004), *Deceptive field experiments of discrimination: are they ethical?*, „Kyklos” 57 (3), 457–470.
- Riach, P.A., & Rich, J. (2006), *An experimental investigation of age discrimination in the French labour market*.
- Riach, P.A., & Rich, J. (2007), *An experimental investigation of age discrimination in the Spanish labour market*.
- Stewart, T.L., Laduke, J.R., Bracht, C., Sweet, B.A., & Gamarel, K.E. (2003), *Do the Eyes Have It? A Program Evaluation of Jane Elliott’s Blue-Eyes/Brown-Eyes Diversity Training Exercise 1*, „Journal of Applied Social Psychology” 33 (9), 1898–1921.
- Teele, D.L. (2014), *Reflections on the ethics of field experiments*. Field Experiments and Their Critics: Essays on the Uses and Abuses of Experimentation in the Social Sciences, 115–140.

West, S.G., & Gunn, S.P. (1978), *Some issues of ethics and social psychology*, „American Psychologist” 33 (1), 30.

Zimbardo, P.G., Cybulko, A., Materska, M., Radzicki, J., Zieliński, M., & Kowalczywska, J. (2015), *Efekt Lucyfera: dlaczego dobrzy ludzie czynią zło?*, Warszawa: Wydawnictwo Naukowe PWN.

Oficjalna strona eksperymentu stanfordzkiego: <https://www.prisonexp.org/> (dostęp 20.01.2020).

Ile trzeba zjeść czekolady, żeby dostać Nobla?

Bisson, J.F., Nejdi, A., Rozan, P., Hidalgo, S., Lalonde, R., & Messaoudi, M. (2008), *Effects of long-term administration of a cocoa polyphenolic extract (Acticoa powder) on cognitive performances in aged rats*, „British Journal of Nutrition” 100 (1), 94–101.

Evans, J.D. (1996), *Straightforward Statistics for the Behavioral Sciences*, Brooks/Cole Publishing: Pacific Grove.

Gerstman B.B., (2013), *Epidemiology kept simple: an introduction to traditional and modern epidemiology*, John Wiley & Sons.

Guilford, J.P. (1954), *Psychometric methods (2nd ed.)*, New York, NY, US: McGraw-Hill.

Maurage, P., Heeren, A., & Pesenti, M. (2013), *Does Chocolate Consumption Really Boost Nobel Award Chances? The Peril of Over-Interpreting Correlations in Health Studies*, 2, „The Journal of Nutrition” 143 (6), 931–933.

Messerli, F.H. (2012), *Chocolate consumption, cognitive function, and Nobel laureates*, „New England Journal of Medicine” 367 (16), 1562–1564.

Nurk, E., Refsum, H., Drevon, C.A., Tell, G.S., Nygaard, H.A., Engedal, K., & Smith, A.D. (2008), *Intake of flavonoid-rich wine, tea, and chocolate by elderly men and women is associated with better cognitive test performance*, „The Journal of Nutrition” 139 (1), 120–127.

Spurious Correlations website (2013), dostępna on-line: <http://www.tylervigen.com/spurious-correlations> (dostęp 20.01.2020).

Werner, S., Malaspina, D., Rabinowitz, J., (2007), *Socioeconomic status at birth is associated with risk of schizophrenia: population-based multilevel study*, „Schizophrenia Bulletin” 18.04.2007, 33(6), 1373–1378.

Czy język angielski powoduje zawały serca, a masło przedłuża życie?

- Artaud-Wild, S.M., Connor, S.L., Sexton, G., & Connor, W.E. (1993), *Differences in coronary mortality can be explained by differences in cholesterol and saturated fat intakes in 40 countries but not in France and Finland. A paradox*, „Circulation” 88 (6), 2771–2779.
- Criqui, M.H., Ringel, B.L., (1994), *Does diet or alcohol explain the French paradox?*, „Lancet” 344, 1719–1723.
- Ducimetiere, P., Cambien, F., Richard, J., Rakotovo, R., & Claude, J. (1980), *Coronary heart disease in middle-aged Frenchmen: comparisons between Paris Prospective Study, Seven Countries Study, and Pooling Project*, „The Lancet” 315 (8182), 1346–1350.
- Durkheim, E. (2006), *Samobójstwo: studium z socjologii*, Warszawa: Oficyna Naukowa.
- Evans, A. (2011), *The French paradox and other ecological fallacies*, „International journal of epidemiology” 40 (6), 1486–1489.
- FAO Food Consumption Nutrients spreadsheet – 2008. Food and Agriculture Organization of the United Nations. Retrieved February 18, 2009.
- Hegsted, D.M., Ausman, L.M., (1988), *Diet, alcohol and coronary heart disease in men*, „The Journal of Nutrition” 118, 1184–1189.
- Jargowsky, P.A. (2005), *The ecological fallacy*, „Encyclopedia of social measurement” 1, 715–722.
- Klatsky, A.L., Armstrong, M.A., Friedman, G.D., (1992), *Alcohol and mortality*, „Annals of internal medicine” 117, 646–654.
- Law, M., Stampfer, M., Barker, D.J.P., Mackenbach, J.P., Wald, N., Rimm, E., & Kunst, A.E. (1999), *Why heart disease mortality is low in France: the time lag explanation*, „British medical journal” 318 (7196), 1471–1480.
- Renaud, S., De Lorgeril, M., (1992), *Wine, alcohol, platelets, and the French paradox for coronary heart disease*, „Lancet” 339, 1523–1526.
- Robinson, W.S. (2009), *Ecological correlations and the behavior of individuals*, „International journal of epidemiology” 38 (2), 337–341.
- Sedgwick, P. (2015), *Understanding the ecological fallacy*, „British medical journal” 351, h4773.
- St Leger, A.S., Cochrane, A.L., Moore, F., (1979), *Factors associated with cardiac mortality in developed countries with particular reference to the consumption of wine*, „Lancet” 1 (8124), 1017–1020.
- The Sentencing project, dostępny on-line: <https://www.sentencingproject.org/issues/incarceration/> (dostęp 20.01.2020).

- Thun, M.J., Peto, R., Lopez, A.D., Monaco, J.H., Henley, J., Heath, C.W., et al., (1997), *Alcohol consumption and mortality among middle-aged and elderly US adults*, „The New England journal of medicine” 24, 1705–1714.
- Westling, T. (2011). *Male organ and economic growth: does size matter?*, dostępny on-line: <https://helda.helsinki.fi/bitstream/handle/10138/27239/maleorga.pdf> (dostęp 20.01.2020).
- United States Sentencing Commission, (2018), *Demographic Differences in Sentencing: An Update to the 2012 Booker Report*, „Federal Sentencing Reporter” 30 (3), 212–229, dostępny on-line: https://www.ussc.gov/sites/default/files/pdf/research-and-publications/research-publications/2017/20171114_Demographics.pdf (dostęp 20.01.2020).

Czego statystyka nauczyła się od martwego łososia?

- Amin, N.P., Martin, S.S., Blaha, M.J., Nasir, K., Blumenthal, R.S., & Michos, E.D. (2014), *Headed in the right direction but at risk for miscalculation: a critical appraisal of the 2013 ACC/AHA risk assessment guidelines*, „Journal of the American College of Cardiology” 63 (25 part A), 2789–2794.
- Austin, P.C., Mamdani, M.M., Juurlink, D.N., & Hux, J.E. (2006), *Testing multiple statistical hypotheses resulted in spurious associations: a study of astrological signs and health*, „Journal of clinical epidemiology” 59 (9), 964–969.
- Bennett, C.M., Miller, M.B., & Wolford, G.L. (2009), *Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: an argument for multiple comparisons correction*, „Neuroimage” 47 (suppl 1), S125.
- Chen, X., Liu, G., Margerie, J., & Reid, M.F. (2008), *A few mistakes in widely used data files for fn configurations calculations*, „Journal of Luminescence” 128 (3), 421–427.
- Ghirlanda, S., Jansson, L., & Enquist, M. (2002), *Chickens prefer beautiful humans*, „Human nature” 13 (3), 383–389.
- Zar, J.H. (1968), *Calculation and miscalculation of the allometric equation as a model in biological data*, „BioScience” 18 (12), 1118–1120.

Dlaczego ufamy szarlatanom?

- Berger, J., & Piotrowska, M. (2013), *Efekt wirusowy w biznesie: dlaczego pewne produkty i usługi zdobywają rynek*, Warszawa: MT Biznes.
- Forer, B.R. (1949), *The fallacy of personal validation: a classroom demonstration of gullibility*, „The Journal of Abnormal and Social Psychology” 44 (1), 118.
- Haase, N., Betsch, C., & Renkewitz, F. (2015), *Source credibility and the biasing effect of narrative information on the perception of vaccination risks*, „Journal of health communication” 20 (8), 920–929.
- Kahneman, D., Tversky A. (1979), *Prospect theory: an analysis of decision under risk*, „Econometrica” 47, 263–292.
- McCandless, D. (2019), *Snake Oil Superfoods. Solid scientific evidence for extra health benefits of certain foods*, dostępny on-line: <https://www.informationisbeautiful.net/visualizations/snake-oil-superfoods/> (dostęp: 20.01.2020).
- National Cancer Institute (2018), *Cancer Treatment. Gerson Therapy*.
- Weisberg, D.S., Keil, F.C., Goodstein, J., Rawson, E., & Gray, J.R. (2008), *The seductive allure of neuroscience explanations*, „Journal of cognitive neuroscience” 20 (3), 470–477.

¹ Do dziś w Muzeum Guinnessa w Dublinie znajduje się tabliczka upamiętniająca Gosseta i jego statystyczne osiągnięcia. Wiem, bo to pozwalało mi organizować studenckie wycieczki do tego muzeum pod pozorem wizyty naukowej, by nie powiedzieć – wizyty w miejscu sakralnym. Co ja rozumiałam jako miejsce powstania najważniejszego testu statystycznego na świecie, a moi studenci jako ten moment, kiedy w cenie muzealnego biletu każdy dostawał piwo.

² Wbrew opinii moich studentów nazwa test t-Studenta nie powstała po to, by upamiętnić tragedię rzeszy studentów, którzy do dziś trzaskają ten test na wszelkich egzaminach. Gosset musiał przyjąć pseudonim, bo kierownictwo Guinnessa nie chciało, by ich konkurencja wiedziała, że opracowali matematyczny wzór na doskonałego browara.

³ Niemniej jeśli znajdzie się tutaj jakiś fanatyk gatunkowej poprawności wśród cytrusów, to zapraszam też na stoisko w Biedronce, niech się bije z ludźmi, którzy mówią: „Daj mi tę pomarańczę”.

⁴ Z tym że trzeba poczynić wiele założeń dotyczących jeża, lasu, a także częstotliwości i trajektorii studenckich spacerów. Ale tak, można to policzyć, bo tym właśnie jest statystyka – nauką, która potrafi oszacować prawie wszystko.

⁵ Miałam kiedyś studenta o tym wdzięcznym (i dźwięcznym) imieniu i mówiłam do niego Parąjk. Przez trzy lata tak mówiłam i dopiero po tym czasie on mi powiedział, że w sumie to jego imię czyta się Poridż, NO SPOKO.

⁶ Ponieważ spodziewam się, że nikt nie będzie przepisywał linku z książki jak jakiś amisz, podpowiadam, że kota można znaleźć w serwisie YouTube po wpisaniu hasła *cat dead owner*.

⁷ O tej rzetelności przeczytacie więcej w rozdziale *Czy naukowcy są jak żony i nigdy się nie mylą?*.

⁸ Zupełnie jak my obecnie. Wiadomo – słynna zasada 5 sekund.

⁹ To jest komplement, ja kocham rodzinę w serniku!

¹⁰ Ganiecie kolegi z łopata nie jest niczym nowym w historii nauki. W 1946 roku dwóch doskonałych filozofów Karl Popper i Ludwig Wittgenstein pokłóciło się o podstawową naturę filozofii. Jedna z anegdot donosi, że nie poprzestali na fechtunku słowem, ale próbowali okładać się rozżarzonymi pogrzebaczami.

¹¹ Pozwalam sobie na ten żart, bo w życiu popełniłam tylko dwa spektakularne błędy: jednym z nich było włożenie jajka do mikrofalówki, drugim – decyzja, by studiować pedagogikę, czyli kierunek, którego celem było zapoznanie studentów z wiedzą jak najbardziej przestarzałą, nieadekwatną i zbędną (czyli zupełnie przeciwną do tej, która mówi, by nie wsadzać tego jajka do mikrofalówki).

¹² Przeprowadzono wiele eksperymentów, które dowiodły, że proces recenzji naukowej dyskryminuje nie tylko koty, lecz także płeć (Helmer et al. 2017) czy... narodowość chińską (Campos-Arceiz et al. 2015).

¹³ Policzyłam, że za jej równowartość można by kupić około 20 512 kilogramów tuńczyka. Czyli całkiem sporo, może to akurat w sam raz, by wasz kotek przestał dzwonić do TOZ-u, że jest głodzony, bo nie jadł nic od pięciu minut.

¹⁴ Jeśli interesuje was ten temat, to koniecznie przeczytajcie książkę Setha Mnookina *Wirus paniki. Historia kontrowersji wokół szczepionek i autyzmu*. To publikacja, która po pierwsze, systematycznie omawia i obala najczęstsze mity związane ze szczepieniami (autyzm, płody po aborcji, „krzyk mózgowy” i wiele, wiele innych) – skąd się wzięły, dlaczego udało im się tak zakorzenić w społecznej świadomości i, przede wszystkim, dlaczego nie są prawdą i nie należy się ich bać. Po drugie, robi coś innego, równie ważnego. Wskazuje wszystkie błędy, które w przeszłości zostały popełnione podczas komunikowania osiągnięć nauki czy radzenia sobie z sytuacjami kryzysowymi i które doprowadziły do tego, że teorie wszystkich szarlatanów tego świata trafiły na podatny i miły im grunt, żyzny na tyle, by kwitły, rozwijały się i wrastały w myślenie i działanie coraz większej grupy ludzi. Ta książka mówi też o tym, w jaki sposób arogancja polityków i naukowców doprowadziła do kryzysu wiary w naukę. I o wszystkich mechanizmach społecznych i błędach poznawczych, które sprawiają, że ruch antyszczepionkowy (a szerzej – ruch antynaukowy) ma się tak dobrze; przytulił się do opinii publicznej i szybko nie odpuści.

¹⁵ Korelacja jest prawdziwa i znajdziecie ją na stronie, która zbiera wszelkie dziwne korelacje, by stale przypominać ludziom, że należy zachowywać pewną dozę ostrożności. Wystarczy, że wpisiecie w wyszukiwarce hasło *tylervigen spurious correlations*.

¹⁶ Widzisz, mamó? Jestem w doborowym towarzystwie!!!

¹⁷ A to tylko początek złych wiadomości – obawiam się, że zdefiniowanie naszej populacji jako „przedszkolaki z Warszawy” wciąż nie jest wystarczające. No bo czy mamy na myśli wszystkie dzieci w wieku przedszkolnym zamieszkałe w Warszawie? Czy też może jedynie te dzieci, które uczęszczają do szkół dla mikroludzi, a nie te, które pozostają w domach z mamami, tatami lub innymi opiekunami?

¹⁸ To taki żart jest. Badania ankietowe na przedszkolakach, a nawet niemowlakach jak najbardziej da się przeprowadzać – wtedy na pytania odpowiadają za nich opiekunowie (w przypadku niemowlaków, bo one są leniwe i nie chce im się samodzielnie wypełniać ankiety) lub pytania i odpowiedzi są czytane na głos przez ankietera (w przypadku przedszkolaków). Bardzo ciekawym przykładem badania, które śledziło losy dzieci od niemowlęctwa aż do nastoletniości, jest irlandzkie badanie longitudinalne *Growing up in Ireland*.

¹⁹ Serio, w irlandzkich siłowniach są maszyny z batonikami. To jedyny powód, dla którego kiedyś chodziłam tam pięć razy w tygodniu.

²⁰ Wyjątkiem jest ten student, który swego czasu codziennie chodził na dokładnie te same moje zajęcia. W poniedziałek przyszedł, we wtorek przyszedł i w środę też przyszedł. Gdy go o to spytałam, wyjaśnił mi, że chodzi ciągle na moje zajęcia, bo codziennie o 17.00 odjeżdża mu autobus do domu i woli czekać na ów autobus w sali, bo na dworze (nie polu!) zimno, z kolei na korytarzach śmierdzi frytkami.

²¹ Autentyk. Przekonała się o tym moja koleżanka, doktorka Anna Oleszkiewicz z Instytutu Psychologii Uniwersytetu Wrocławskiego, która opowiadała mi kiedyś, że kazała studentom na zajęciach zapisać plik poprzez kliknięcie na dyskietkę, ale studenci nie mogli wykonać polecenia, bo nie wiedzieli, co to jest dyskietka. Śmiałyśmy się z tego przez całą drogę do sklepu z katafalkami, bo bardzo wyraźnie wtedy poczułyśmy, że chyba już czas się w takowy zaopatrzyć.

²² Ponadto owce mają własne upodobania, jeśli chodzi o pozostałe owce, to znaczy w sensie poszukiwania wełnianej sympatii – fizjonomia niektórych owiec podoba im się bardziej niż innych, bo z niektórymi to, no dobra, od biedy pójdą na frytki, ale już na studniówkę ich nie zaproszą (Kendrick et al. 1995, Kendrick et al. 1996).

²³ Na szczęście wszyscy czytaliśmy poprzedni rozdział, więc wiemy już, dlaczego takie publikowanie ankiet na Facebooku jest zupełnie bez sensu.

²⁴ To jest hipotetyczny przykład, proszę mi nie przysyłać zgłoszeń ani tym bardziej swoich dzieci do udziału w tym eksperymencie.

²⁵ Więcej o eksperymentach, quasi-eksperymentach i eksperymentach terenowych piszę w rozdziale *Czy wszyscy jesteście rasistami?*.

²⁶ Mam tu na myśli Daniela Kahnemana oraz Amosa Tversky'ego i ich badania eksperymentalne dotyczące procesu podejmowania decyzji – prosili oni badanych o wybranie jednej z dwóch możliwości (na przykład: zbliża się epidemia, która może zabić 600 osób; masz do wyboru dwa programy; który wybierzesz, by uratować ludzkość?). W dużym uproszczeniu wykazali, że to, w jaki sposób sformułowano komunikat, wpływa na nasze decyzje, nawet gdy między opcjami do wyboru nie ma tak naprawdę żadnej różnicy. Najważniejsze wnioski zostały zebrane w książce *Pułapki myślenia. O myśleniu szybkim i wolnym*. Uważam, że absolutnie każdy powinien ją przeczytać. ALE NIE TERAZ, dopiero jak skończycie TĘ książkę!!!

²⁷ Co ciekawe, pojechał tam ze swoim najlepszym przyjacielem, który jednocześnie był fotografem i rysownikiem. Aha, ten przyjaciel się nazywał – Witkacy. Ale Witkacy towarzyszył Malinowskiemu jedynie na początku podróży, bo jak tylko dowiedział się o wybuchu wojny, wrócił do kraju i zaciągnął się do carskiej armii.

²⁸ Wiedzą o tym doskonale moi rodzice, którzy mając trójkę dzieci, równo dzielili między nie swoje rodzicielskie uczucia, nawet w przypadku porażek wychowawczych w postaci zapalonego papierosa, późnych powrotów po jednym piwku za dużo albo tego wspomnianego już momentu, kiedy oznajmiłam im, że chcę studiować pedagogikę.

²⁹ Sudhir Venkatesh, który studiował na uniwersytecie w Chicago, miał zresztą bardzo dobre wzorce. Na tej samej uczelni w 1892 roku powstał pierwszy wydział socjologii na świecie! Tam też rozpoczął się nurt w socjologii zwany szkołą chicagowską, której przedstawiciele podkreślali konieczność badań empirycznych, również terenowych, a szczególnie interesowały ich socjologia miasta i socjologia dewiacji. Znajdziemy tam także polski akcent, bo jedną z ważniejszych prac, które powstały w ramach tego nurtu, była książka Floriana Znanieckiego i Williama Thomasa o jakże nęcącym tytule *Chłop polski w Europie i Ameryce*. Pomyślcie o tym w ten sposób – to był jeden z naszych pierwszych towarów eksportowych: dobra socjologia. Obecnie został niestety wypchnięty przez jabłka, pieczarki, soki zagęszczone, drób i przetwory mleczne.

³⁰ W oryginale: *The venom injected by the bite of the adult female spider, Latrodectus mactans, is dangerously poisonous for man.*

³¹ Szkoda, że nikt nie powiedział tego innemu badaczowi, Michaelowi Smithowi, który w 2014 roku dał się użądlić pszczołom w 25 różnych części ciała, żeby sprawdzić, które z tych ugryzień będą najbardziej bolesne. Na wszelki wypadek podaję wyniki jego badań, byście nie musieli sami

tego próbować: najgorsze są ugryzienia w nozdrza, górną wargę i penisa. W ostatnim przypadku brak jest żeńskiego ekwiwalentu, więc gdyby ktoś szukał pomysłu na doktorat, to jest luka w badaniach.

³² Dobra, wiem, że wszyscy chcecie to teraz zobaczyć! Owcę miziącą prezydenta chrapkami zobaczycie na stronie National Geographic po wpisaniu frazy: *sheep recognize celebrity faces*.

³³ Co w sumie dowodzi tego, że owce bywają bystrzejsze od Janin. Ja egzaminowałam kiedyś studenta, którego zdecydowanie skądś znałam, i w ogóle nie mogłam sobie przypomnieć skąd. Cały egzamin myślałam, on pisał, ja myślałam, on napisał, ja nie wymyśliłam. Porozmawialiśmy chwilę, poszedł, a ja, jako że mam zapłon jak PKS na trasie Hyżne Nowa Wieś – Nieborów, to już po niecałych trzech godzinach skminiłam, skąd go znam. I do dziś uważam, że król Joffrey z *Gry o tron* nie może mieć do mnie pretensji – jeśli przychodzi na egzamin bez korony, to niech się później nie dziwi, że go nie poznaję.

³⁴ Badanie przeprowadzone przez sztokholmskich badaczy wyglądało następująco: kurczaki zostały wytrenowane w taki sposób, by poprzez dziobanie wskazywać zdjęcie, które podoba im się najbardziej. Następnie pokazywano im zdjęcia ludzi o różnej atrakcyjności fizycznej i proszono każdego kurczaka o opinię. Jednocześnie identyczne zadanie dostała grupa studentów – w sensie mieli oceniać atrakcyjność zdjęć przedstawiających różnych atrakcyjnych ludzi (nie że zdjęcia mniej lub bardziej ponętnych kurczaków). Wyniki wypadły szalenie korzystnie dla niektórych osób ze zdjęć, albowiem okazało się, że zarówno studenci, jak i kurczaki dzielą to samo umiłowanie do piękna i chętniej dziobią (to kurczaki) lub zakreślają (to studenci) dokładnie te same zdjęcia.

³⁵ Ich pytanie badawcze brzmiało zupełnie inaczej. Chcieli wiedzieć, jakie preferencje dotyczące ludzkich twarzy mają kurczaki i które uważają za ładne, niemniej ja z upodobaniem odnalazłam w tym badaniu komercyjny potencjał.

³⁶ Było to dość okrutne badanie, w którym nikt z nas nie chciałby wziąć udziału, zwłaszcza po świętach – badacze sadzali kurczaki na grzędzie-wadze, a chwilę później próbowali zrobić to znowu i sprawdzali, na ile kurczaki były chętne ponownie wejść na wagę.

³⁷ Albo nie pobieramy, gdy przeprowadzamy cenzus, czyli badanie na całej populacji.

³⁸ Średnia ważona to rodzaj średniej, w której poszczególnym elementom przypisywane są różne znaczenia, czyli wagi. Elementy o większej wadze mają większy wpływ na średnią niż pozostałe.

³⁹ Zaznaczyć chciałam, że co do takiego traktowania liczby nóg nie wszystkie statystyczne świry są zgodne. Są i tacy, którzy uważają, że w tym wypadku liczba nóg jest zmienną ilościową, i tacy, według których owa liczba nóg może zostać potraktowana jako zmienna nominalna o dwóch poziomach (dwa albo cztery). W gruncie rzeczy na tym polega piękno nauki, że nie zawsze się ze sobą zgadzamy, ale to jest zupełnie OK, wcale nie musimy się z tego powodu okładać surowym okoniem po pysku.

⁴⁰ A także bawi, cieszy i jest źródłem niezliczonych anegdotek przy kolacji wigilijnej.

⁴¹ Jest pewną ironią kolorystyczną, że te szare komórki są również niezbędne, by zrozumieć, że to wcale nie jest tak, że średnio każda zebra na świecie jest szara, he, he, he.

⁴² Jedynym przypadkiem, w którym procenty na wykresie mogą w sumie wynosić ponad 100%, jest ten, kiedy wykres dotyczy pytania wielokrotnego wyboru, czyli respondenci mogli zaznaczać

więcej niż jedną odpowiedź.

⁴³ Oczywiście oprócz tego pytania, które zadają zdecydowanie najczęściej, czyli: ile można mieć nieobecności?

⁴⁴ Szczegółowy ranking wizualizacji od najlepszych do najgorszych: wspólna skala, identyczna skala, porównanie długości, kątów i nachylenia, powierzchnia, objętość i na końcu – zacienienie, kolor, wysycenie.

⁴⁵ W tej drugiej sytuacji sprawę dodatkowo utrudnia to, że wiele osób jest bardzo blisko zapadnięcia w śpiączkę cukrzycową po tym, jak właśnie opróżnili talerz z piramidy Cheopsa zrobionej z ciastek, bo przecież koszty opłaty konferencyjnej same się nie zwrócą.

⁴⁶ W oryginale: *Information is beautiful*. Taki tytuł mają TED Talk Davida McCandlessa, jego książka i blog, na którym pokazuje najciekawsze sposoby wizualizacji danych, często w formie interaktywnej. Jednym z moich ulubionych wykresów jest na przykład ten przedstawiający, jak zmieniało się w czasie podejście do kluczowych teorii naukowych, który znajdziecie poprzez wyszukanie na jego stronie haseł: *mavericks and heretics*.

⁴⁷ Innym źródłem doskonałych interaktywnych wizualizacji jest strona fundacji Gap Minder założona przez wspomnianego już wcześniej Hansa Roslinga i jego drużynę: <https://www.gapminder.org>.

⁴⁸ Jest twórcą między innymi współczynnika kłamstwa, współczynnika dane/atrament czy współczynnika perspektywy miniwykresu. Jak widzicie, ten człowiek kochał wizualizację danych prawie tak bardzo jak słowo „współczynnik”.

⁴⁹ Najpiękniejsze okazy takich wykresów znajdziecie między innymi na mojej ulubionej stronie na FB: Analityk płakał, jak raportował.

⁵⁰ Pomysł ten został wcześniej wykorzystany w infografice opublikowanej w „South China Morning”. Niemniej tam został zaimplementowany poprawnie i w sposób, który nikogo nie konfundował. Znajdziecie go po wpisaniu w wyszukiwarce *Iraq's Bloody Toll* (autorem wykresu jest Simon Scarr).

⁵¹ Choć jednocześnie nie należy zapominać o tym, że przeprowadzenie takiego badania wymagałoby określonych kompetencji, na przykład mój kolega Kamil nijak nie mógłby takim pomiarem zarządzać, albowiem on twierdzi, że nie ma pojęcia, co to jest za zwierzę „oposa”.

⁵² Niemniej przypominam, że wybór ten zawsze powinien być uzasadniony i poprzedzony dokładnym i krytycznym przeglądem literatury. W sensie literatury związanej z tematem, a nie że *Anny Kareniny*.

⁵³ Generalnie to ja uważam, że biegać należy tylko wtedy, kiedy ktoś nas goni.

⁵⁴ Według badań Jana Błuszkowskiego (2006) jesteśmy dla siebie dość surowi – przypisujemy sobie narodowo następujące wady: nadużywanie alkoholu, rozrzutność, lenistwo i brak dyscypliny. Ja jestem odpowiedzialna tylko za 90% tych wskazań, nie ma za co. Niemniej tak naprawdę nie znalazłam badań, które potwierdzałyby tę obiegową i niezbyt sympatyczną opinię o naszym narodzie, jakoby cudza krzywda cieszyła nas najbardziej.

⁵⁵ Kraje OECD to 36 wysoko rozwiniętych i demokratycznych państw zrzeszonych na mocy konwencji podpisanej 14 grudnia 1960 roku (początkowo przez 20 państw). Są to kraje nie tylko

europeskie – znajdziemy tam również Kanadę, Stany Zjednoczone, Chile, Australię lub Koreę Południową.

⁵⁶ Jak dokładnie poradziłyśmy sobie w poszczególnych sferach, można sprawdzić na stronie <http://www.oecdbetterlifeindex.org/> po wyszukaniu Polski.

⁵⁷ We mnie osobiście to pytanie powoduje pewien dysonans, bo w sumie to lubię myśleć, że jestem na dość wysokim stopniu tej metaforycznej drabiny szczęśliwości, ale z drugiej strony nie mam żadnych wątpliwości, że gdybym weszła tyle szczebli do góry, to byłabym raczej bardzo nieszczęśliwym człowiekiem (zwłaszcza że na górze jakichkolwiek drabin rzadko występuje Wars).

⁵⁸ Uproszczony wzór na światowy wskaźnik szczęścia wygląda tak: odczuwany dobrostan pomnożony przez oczekiwaną długość życia i to wszystko podzielone przez ślad ekologiczny. Dokładny wzór jest trochę bardziej skomplikowany i nie będziemy się tutaj zajmować jego liczeniem, zwłaszcza że od lat z powodzeniem zajmuje się tym Eurostat, czyli świątynia europejskich statystyk, portal otwartych danych dotyczących UE.

⁵⁹ To zdanie to wyraz najwyższego optymizmu pedagoga. Śmieję się za każdym razem, gdy je wypowiadam na zajęciach. Ej, pamiętacie z poprzednich zajęć... – pytam studentów i już w połowie tego zdania dociera do mnie naiwność tego pytania, dajcie spokój, wszyscy byliśmy na studiach i wiemy, że tam człowiek jest tak zajęty zjeżdżaniem ze schodów w akademiku na sankach zrobionych z pudełka po pizzy, że w ogóle nie ma czasu czeokolwiek zapamiętywać.

⁶⁰ Biorąc pod uwagę, że sporo czytających mnie osób pracuje na uczelni, doszłam do wniosku, że warto wyjaśnić tym miłym ludziom, co to są pieniądze. Otóż już tłumaczę: to jest takie coś metalowe lub papierowe, co dostaje się za wykonaną pracę i można to wymieniać na różne dobra typu bułka, bilet autobusowy, mop parowy. Chyba że pracujecie na uczelni publicznej, to nie można, bo już dawno przekonano was, że najlepszą walutą jest satysfakcja, a najlepszym śniadaniem – kanapka z gruzem i posypką ze sproszkowanej cegły.

⁶¹ To nie jest prawdziwe stowarzyszenie, proszę się nie jarać i nie pędzić, by się zapisać.

⁶² Ja rozumiem, że łabędzie nie pływają po morzu, ale tym bardziej czyni to tę metaforę tragiczną.

⁶³ Błagam, nie kopcie gołębi na ulicy. To nie ich wina, że tak irytująco patrzą.

⁶⁴ Ciekawostka: jeśli chodzi o rapowanie, to mamy już w nauce pewne osiągnięcia na tym polu. Kojarzycie Wielki Zderzacz Hadronów? Przygotowania do włączenia tego urządzenia przypominały trochę lekcje tańca dla nowożeńców – kilka osób w sali świetnie się bawi, a cała reszta nie ma pojęcia, co się dzieje, o co chodzi i za jakie grzechy. By trochę sprawę ułatwić, kilkoro pracujących w CERN-ie naukowców napisało, wykonało i nagrało rapowany utwór, który wyjaśniał, czym się tam zajmują. Wspaniale to jest, sami zobaczcie: <https://youtu.be/j50ZssEojtM> (jeśli klikacie w link i wam nie działa, bo na przykład macie w rękach drukowaną książkę, to wystarczy że wpiszeć w wyszukiwarce *Large Hadron Rap*).

⁶⁵ Oczywiście ta metafora nie dotyczy pierogów, bo one mają wyłącznie zalety.

⁶⁶ Jeśli ktoś jest psychofanem błędów i problemów związanych z badaniami ankietowymi, to najpewniej ucieszy go prezent w postaci książki *Survey errors and survey costs* autorstwa Roberta M.

Grovesa. Tam wszystkie te błędy są bardzo zgrabnie opisane wraz z podziałem na poszczególne dyscypliny.

⁶⁷ Wyobraźcie sobie na przykład, że ktoś najpierw prosi was o to, by wymienić jak najwięcej problemów społecznych i ekonomicznych, które targają naszym krajem, a w kolejnym punkcie pyta, czy myślicie, że nasz prezydent robi dobrą robotę.

⁶⁸ Dla przykładu w czasie okołowyborczym musimy uważnie sprawdzać, czy sondaże pytały o to, jakiego kandydata ktoś POPIERA czy też na kogo GŁOSOWAŁ – w gruncie rzeczy są to dwa różne pytania i mogą prowadzić do dwóch różnych odpowiedzi!

⁶⁹ Zbadali to między innymi Williams i Bargh (2008).

⁷⁰ Jak wynika z ankiety przeprowadzonej przez VoucherCloud i cytowanej przez „Los Angeles Times”. Wrócimy do tego badania w kolejnym rozdziale.

⁷¹ To częsty błąd obecny we wszelkich doniesieniach o tym i wielu innych badaniach, jako że zostało ono przeprowadzone na niereprezentatywnej próbie użytkowników portalu, to nie jest prawdą, że 11% AMERYKANÓW boi się HTML przenoszonego drogą płciową – boi się tego co dziesiąty RESPONDENT, KTÓRY WZIAŁ UDZIAŁ W BADANIU.

⁷² Ten błąd nazywamy *self-selection bias*; polega on na tym, że osoby, które na ochotnika biorą udział w badaniu różnią się znacząco cechami demograficznymi od pozostałych. Zresztą dotyczy to nie tylko osób badanych – podobny błąd ma miejsce, gdy analizujemy, kto wystawia naszym produktom lub usługom oceny i opinie w internecie. Częściej robią to osoby, które mają ekstremalne opinie o produkcie – pozytywne lub negatywne (zobacz Pavlou i Zhang 2017 czy Li i Hitt 2007).

⁷³ Są to tak zwane pytania sugerujące. Przykładami są nie tylko pytania typu „Czy podoba Ci się ta książka i dlaczego tak?”, lecz także pytania typu „WIĘKSZOŚĆ POLAKÓW JEST PRZECIWKO rodzinom w serniku; czy Pan/Pani również jest PRZECIW?”. W tym wypadku oba podkreślone wyrażenia mogą w jakiś sposób wpłynąć na odpowiedź respondenta.

⁷⁴ Tak naprawdę takie badania nad skalą, pytaniami, długością kwestionariusza czy nawet formą wstępu mają sporo sensu i dość proste uzasadnienie: walczymy o jak najwyższą wypełnialność ankiety. Jest to problem zwłaszcza w przypadku kwestionariuszy typu CAWI, gdy nie ma ankietera, a respondentów łatwo może rozproszyć zdjęcie kota z chlebem na głowie albo gołąb za oknem, który wygląda jak Elvis Presley. Jeśli ktoś jest zainteresowany tym tematem albo też brakuje mu ciekawych tematów do poruszenia na randce, to sporo (i wspaniale) piszą o tym Mick P. Couper czy Don A. Dillman. W bibliografii znajdziecie kilka ich publikacyjnych diamentów.

⁷⁵ Oba te zdarzenia nazywamy *recall bias*, niemniej warto pamiętać, że mimo wspólnej nazwy mają zupełnie inne przyczyny. Czasem odpowiedź niezgodna z rzeczywistością wynika z tego, że zbyt ciężko obciążamy pamięć respondenta. Czasem jednak w epidemiologii obserwuje się inne zjawisko, mianowicie osoby chore lub bliscy osób chorych częściej raportują styczność z potencjalnymi czynnikami ryzyka, szukając wyjaśnień, dlaczego przydarzyły im się złe rzeczy (zobacz na przykład Infante-Rivard i Jacques 2000).

⁷⁶ W sumie nie jestem pewna, czy taka obecność dzieci mogłaby wpłynąć na odpowiedzi w sposób pozytywny, czy negatywny. To pewnie zależy od tego, czy dzieci akurat rzucałyby w przechodniów kasztanami, czy też rozdawały im kompot pozostały z obiadu.

⁷⁷ W sumie już sam tytuł tego artykułu jest spoilerem, co wskazuje na to, że ów francuski zespół może i na badaniach naukowych się znał, ale na budowaniu suspense niekoniecznie. Najpewniej pisaliby najgorsze kryminały świata.

⁷⁸ Oprócz tych trzech wersji badano również reakcję szczurów na ekwiwalentne stężenie sacharozy, czyli naturalnego cukru, ich wybór między kolejnymi dawkami kokainy a słodzikiem, a także zmienność ich preferencji w czasie. Sprawdzano też, jak trudność zdobycia nagrody wpływa na jej wybór. Oryginalny artykuł jest dostępny on-line (link znajdziecie w bibliografii), więc kto ciekawy, może się zapoznać z wszystkimi szczegółami.

⁷⁹ Osoby bardziej zainteresowane tematem najpewniej zainteresują się również artykułem *Chocolate. Food or drug?* (autorstwa Kristen Bruinsma i Douglasa L. Tarena). Próbowano w nim wyjaśnić, dlaczego jest tak, że w prawie połowie przypadków, gdy mamy ochotę coś przegryźć, to sięgamy po czekoladę (a w 80% – po przekąskę, która zawiera cukier). W gruncie rzeczy wyniki są dla nas korzystne w tym sensie, że dowodzą, iż niepohamowana ochota na czekoladę jest zjawiskiem jak najbardziej prawdziwym i udowodnionym. Najlepiej więc wydrukować sobie ten artykuł i okazywać go wszystkim w cukierni, gdy będą oceniać nasz zakup kolejnego kawałka brownie.

⁸⁰ Czasem nie mamy możliwości zastosowania losowego doboru badanych do grup badawczych. W takim wypadku mamy do czynienia z badaniem quasi-eksperymentalnym.

⁸¹ Pak to takie czarne, szkliste, bezpostaciowe coś (to wersja dla ludzi, którzy nie słuchali na lekcjach chemii, bo byli zajęci podrabianiem zwolnień: „Proszę usprawiedliwić nieobecność mojego syna. Z poważaniem, Mama”). Wersja dla zaawansowanych: to ciecz o bardzo dużej lepkości, która jest pozostałością po destylacji smoły pogazowej.

⁸² Po jego śmierci na chwilę o eksperymencie zapomniano, niemniej w latach 60. wznowił go profesor John Mainstone. Opiekował się nim przez kolejne 52 lata, aż do śmierci. Otrzymał za niego Nagrodę Ig Nobla (wraz z Thomasem Parnellem, z tym że ten drugi pośmiertnie).

⁸³ Dotychczas nikt nie był takim szczęściarzem, by osobiście zaobserwować spadającą kroplę paku, nawet twórca eksperymentu.

⁸⁴ W przypadku zajęć na dworze kluczowe jest, by wybrać ich lokalizację wystarczająco daleko od wydziału technicznego. To na wypadek, gdyby koledzy ścisłowcy zechcieli w porze lunchu rzucić koszule w kratę i się trochę poopalać, bo ja potem nie mam czasu tłumaczyć moim studentom, że ci biali jak śnieg ludzie, co to odbijają światło słoneczne skuteczniej od samego Księżyca, to wcale nie są chorzy, tylko z informatyki.

⁸⁵ Rudy wraz z kolegą wrócili do grupy po blisko dziesięciu minutach i podzielili się z wszystkimi jakże szalenie istotną informacją, jakoby prawie dogonili kota, ale im uciekł, gdy zatrzymali się na chwilę popatrzeć, jak mewa wymiotuje myszą.

⁸⁶ Na wszelki wypadek od razu zaznaczam, że oba zawody są tak samo wartościowe. W sumie to wykładanie chemii w supermarkecie bywa często znacznie ważniejsze od wykładania na uniwersytecie, bo tam to przynajmniej można realnie dokonać jakiejś zmiany, a uczenie innych często przypomina raczej zamiatanie kałuży grabiami.

⁸⁷ Dokładniej rzecz ujmując, badacze skupili się na dwóch miastach – Bostonie i Chicago, i na czterech obszarach zatrudnienia: sprzedaż, administracja, prace biurowe i obsługa klienta.

Przygotowali również życiorysy w dwóch wersjach: w wysokiej jakości oraz takie wyglądające jak napisane przez geś, której ortografię sprawdzał jeź.

⁸⁸ To, czy imię jest „typowo afroamerykańskie”, czy „typowo białe”, zostało ustalone na podstawie oficjalnych statystyk imion, a potem dodatkowo zweryfikowane ankietą, w której pytano, z jakim pochodzeniem kojarzy się badanym dane imię.

⁸⁹ Niepokojąco podobnym wnioskiem zakończyło się inne, bardzo podobne badanie, tym razem przeprowadzone w Nowym Jorku (Pager et al. 2009). Tam również aplikanci o białym kolorze skóry dostali o połowę więcej zaproszeń na rozmowę lub ofert pracy niż osoby czarnoskóre i pochodzenia latynoskiego o identycznych kwalifikacjach. Inne, podobne badania również nie różniły się wnioskami (Kang et al. 2016), zresztą nie tylko na amerykańskim rynku pracy, lecz także w Niemczech (Kaas i Manger 2012) czy Irlandii (McGinnity i Lunn 2011).

⁹⁰ Podobnie sytuacja ma się z kobietami, które choć mają takie same kompetencje w CV, to rzadziej niż mężczyźni są zapraszane na rozmowy kwalifikacyjne (Gonzalez M.J. et al. 2019)

⁹¹ Warto tu jednak wspomnieć, że wokół metodologii eksperymentu stanfordzkiego narastają kolejne pytania (Le Texier 2019, Blum 2018). Co więcej, nie należy tego badania nazywać eksperymentem przynajmniej z dwóch powodów: po pierwsze – brak grupy kontrolnej, po drugie, nie wiemy, jakim czynnikiem manipulowano. Na to drugie pytanie nie potrafi odpowiedzieć sam Zimbardo.

⁹² W 1971 roku grupa psychologów z Uniwersytetu Stanforda wymyśliła sobie, że warto sprawdzić, jak ludzie będą zachowywać się w warunkach symulujących życie więzienne. Naukowcy doskonale zdawali sobie sprawę z tego, że nie ma najmniejszego sensu pytać o to badanych, bo takie odpowiedzi miałyby niewiele wspólnego z rzeczywistością. Philip Zimbardo ze swoją drużyną postanowili więc podejść do tematu inaczej i zrobić to, co każdy z nas ma czasem ochotę uczynić w robocie, to znaczy wybudowali więzienie w piwnicy wydziału psychologii. Co rozczarowujące, wcale nie chcieli tam wrzucać ludzi, którzy pomylili korelację z kauzacją albo – co gorsza – twierdzili, że statystycznie to, he, he, he, każda zebra jest szara. Pomysł mieli inny: zrekrutowali 18 studentów, z czego dziewięcioro zostało wytypowanych na więźniów, a dziewięcioro na strażników, którzy pracowali trójkami podczas ośmiogodzinnych zmian. Celem było zaobserwowanie, jak osoby badane – zarówno „strażnicy”, jak i „więźniowie” – będą zachowywać się w tak ekstremalnych warunkach i w jaki sposób sytuacja społeczna wpłynie na ich zachowania. Nie będę nikomu zdradzać zakończenia, w bibliografii znajdziecie odniesienia do oficjalnej strony eksperymentu i do książki głównego pomysłodawcy owego badania.

⁹³ Kontekst tego badania był trochę szerszy – zostało ono przeprowadzone w 1963 roku, kiedy pamięć o nazistowskich zbrodniach II wojny światowej była jeszcze bardzo świeża. Milgram chciał się dowiedzieć, co doprowadziło zwykłych ludzi do zbrodni dokonywanych między innymi w obozach koncentracyjnych. Jego początkowa teza była taka, że Niemcy jako naród są bardziej od innych narodowości skłonni do ślepego posłuszeństwa wobec autorytetów.

⁹⁴ Technika, w której prawdziwy cel badania jest ukrywany, by ta wiedza w żaden sposób nie wpłynęła na decyzję i zachowania badanych, jest nazywana PROCEDURĄ MASKUJĄCĄ lub DECEPCYJNĄ.

⁹⁵ Co ciekawe, w 2015 roku badanie Milgrama powtórzyła grupa polskich naukowców kierowana przez profesora Tomasza Grzyba (Doliński et al. 2017). Uzyskali wyniki podobne do oryginalnego eksperymentu.

⁹⁶ Gdy po raz pierwszy wypowiedziałam to zdanie do studenta, to totalnie poczułam się jak pełnoprawny człowiek-starość, no bo dajcie spokój, to są takie sformułowania, których z reguły używają STARZY LUDZIE. Na przykład przypadkiem przy obiedzie mówisz mężowi, że ziemniaki może zostawić, ale mięsko niech zje, albo przypominasz swojemu współlokatorowi, że dom to nie hotel. A to nie był koniec. Potem było już tylko gorzej – wtedy kiedy to upuściłam na ziemię swój nowy telefon służbowy (Nokia 3310, rocznik 2001), co szalenie zaniepokoiło mojego studenta (rocznik 1996), który spytał mnie, czy nie boję się, że przez ten upadek telefon się popsuł, a ja mu powiedziałam, że o to bym się nie martwiła, martwiłabym się raczej o to, że właśnie wywołałam wstrząsy sejsmiczne w Dublinie i okolicach, a on mi powiedział, że rzeczywiście ten telefon wygląda na dość duży i stary, jakiego on w ogóle używa oprogramowania, najstarszego Androida?! Pamięci wszystkich, którzy pamiętają erę B.A. (Before Android).

⁹⁷ Nie mam jasnej odpowiedzi na to, czy faktycznie uświęca. Na swoich zajęciach przeznaczyłam cztery godziny na próbę odpowiedzi na to pytanie, z których to czterech godzin jedynie 50 minut spędziłam, czekając na moich spóźnionych studentów, którzy nie dotarli do sali na czas, bo – jak mi wytłumaczyli – byli zajęci zastanawianiem się, czy ten ptak, co leży przed budynkiem, to na pewno jest martwy. Na wszelki wypadek postanowiłam zignorować tę sytuację, głównie z takiego powodu, że nie potrafiłam się zdecydować, co niepokoi mnie bardziej – niezdrowa fascynacja moich studentów nieżywymi zwierzętami czy też fakt, że potrzebują aż 50 minut na to, żeby dojść do wniosku, że ptak pozbawiony głowy jest naprawdę martwy.

⁹⁸ „Eksperyment” w tym wypadku (a także w przypadku eksperymentów Milgrama i Zimbardo) jest słowem, którego używam w jego potocznym rozumieniu, bo nie, nie spełniał on kryteriów eksperymentu badawczego.

⁹⁹ To była podstawówka, a więc ten czas, kiedy przebywanie daleko od pani jeszcze uważaliśmy za karę. Większy błąd popełniliśmy tylko w przedszkolu, gdy za karę uważaliśmy leżakowanie.

¹⁰⁰ Pierwszy film ma tytuł *The eye of the storm*, a drugi – *A class divided*. Krótkie fragmenty obu filmów, a także liczne wywiady z Jane Elliott są dostępne na YouTube.

¹⁰¹ Ej, a czy było tak, że gdy spojrzeliście na ten wykres, to natychmiast zaczęliście szukać własnego kraju? To jest zjawisko, które ja osobiście nazywam patriotyzmem danych, i to jest co prawda zagadnienie nieudowodnione naukowo, ale uważam, że jest tylko kwestią czasu, kiedy świat naukowy doceni moje odkrycie, i jeśli będzie trzeba, to będę codziennie obtaczać się czekoladą, żeby zdobyć za swoją teorię odpowiednią nagrodę.

¹⁰² Dla zainteresowanych: oprócz współczynnika korelacji liniowej Pearsona wyróżniamy również m.in. współczynnik korelacji rang Spearmana czy współczynnik zbieżności Czuprowa. To, który współczynnik korelacji wybierzemy, zależy między innymi od wielkości naszej próby czy rodzaju skali pomiarowej, na jakiej jest mierzona zmienna.

¹⁰³ W oryginale: *Chocolate consumption enhances cognitive function, which is a sine qua non for winning the Nobel Prize, and it closely correlates with the number of Nobel laureates in each country*

(Messerli 2012: 1564).

¹⁰⁴ A wiecie, że wspomniany wcześniej Francis Galton, jeden z moich ulubionych ludzi na świecie, który jako pierwszy opisał model regresji i zastosował go w naukach społecznych, był kuzynem Karola Darwina? Mój Boże, wyobraźcie sobie te wigilie, co za trauma dla pozostałych członków ich rodzin! „Nie, Kaziu, w tym roku nie dostaniesz prezentu, bo nie stworzyłeś żadnej teorii naukowej i nie chcę słyszeć tłumaczeń, że masz dopiero sześć lat!”.

¹⁰⁵ Dane zostały dostarczone przez The Organisation for Economic Cooperation and Development (OECD).

¹⁰⁶ To ostatnie jest w ogóle bardzo ciekawym zagadnieniem. Według danych Organizacji Narodów Zjednoczonych w latach 2005–2006 absolutnym liderem spożycia kalorycznego były... nie, nie Stany Zjednoczone. Te zostały zdetronizowane przez Austrię, która w tamtych latach spożywała średnio 3800 kcal na głowę (a raczej żołądek) każdego obywatela. Francja była niewiele gorsza – 3530 kcal (12. miejsce w rankingu), a jeśli chodzi o narodowe radości, to nasze średnie spożycie wynosiło 3410 kcal, z czego ja jestem odpowiedzialna tylko za połowę wyniku.

¹⁰⁷ To zjawisko to tzw. *racial bias*. Obecnie w Stanach Zjednoczonych ponad 60% osób odbywających karę więzienia to ludzie o innym kolorze skóry niż biały (*The sentencing project* 2018). Ponadto w latach 2011–2016 różnica w długości otrzymanych wyroków przez czarnoskórych i białych mężczyzn, którzy popełnili te same przestępstwa, wynosiła aż 19% na niekorzyść czarnoskórych więźniów (*United States Sentencing Commission* 2012).

¹⁰⁸ W swoim słynnym badaniu Emil Durkheim po raz pierwszy w historii socjologii przeanalizował ogromne ilości statystyk w celu zbadania zjawiska samobójstwa. Na podstawie tych analiz wyróżnił cztery typy samobójstw: samobójstwo altruistyczne, egoistyczne, fatalistyczne i anomiczne. Mimo iż niektóre wnioski z tej publikacji (*Samobójstwo. Studium z socjologii*, wyd. 1897) są obarczone błędem ekologicznym rozumowania, wciąż jest to bardzo ważna i kluczowa pozycja w historii nauki.

¹⁰⁹ Autorstwo tej frazy najczęściej przypisuje się Markowi Twainowi, choć (podobno) cytował on tylko w swojej autobiografii brytyjskiego premiera Benjamina Disraeliego.

¹¹⁰ Ej, a słyszeliście ten huk? To najpewniej jakiś neurobadacz właśnie zemdłał po przeczytaniu mojego brawurowego opisu funkcjonalnego rezonansu magnetycznego. W trosce o dobro bohaterów nauki wszystkich prymusów żądnych dokładnego opisu badania Bennetta i łososa zapraszam do artykułu źródłowego.

¹¹¹ Nie wiem, czy wiecie, ale kiedyś w Tychach ogłoszono konkurs na nazwę nowego parku wodnego. Konkurs trwał trzy miesiące, przyjęto ponad tysiąc zgłoszeń, trzyosobowa komisja obradowała kilka dni, a w końcu zwyciężyła nazwa... Park Wodny Tychy.

¹¹² Niegdyś tłumaczyłam błędy pierwszego i drugiego rodzaju na przykładzie testów ciążowych, które się mylą, ale okazało się, że ten przykład budzi popłoch wśród moich studentów, dlatego szybko przerzuciłam się na wirusy komputerowe, które zdają się wywoływać znacznie mniejsze emocje i nie powodują stanów lękowych.

¹¹³ O wynikach swoich klawych badań Ivan Schwab wspaniale opowiada w wykładzie TEDx *Evolutionary Design in Sight: Ivan Schwab at TEDx Orange Coast*, który jest dostępny na YouTube.

¹¹⁴ W oryginale: *to honor achievements that first make people laugh, and then make them think.*

¹¹⁵ O tym pisał również Peter Austin, który przeprowadził badanie eksperymentalne, w którym celowo nie zastosował odpowiedniej korekty dla tej metody analizy. Skutek był taki, że jego wyniki badań jasno wskazywały na zależności między znakiem zodiaku a zdrowiem. Okazało się między innymi, że ludzie urodzeni pod znakiem Lwa mają większe prawdopodobieństwo krwotoku z układu pokarmowego, a Strzelce – złamania kości ramiennej. Inne znaki zodiaku mogły spać spokojnie, a w sumie po wprowadzeniu odpowiedniej korekty do obliczeń Lwy i Strzelce też, bo wyniki okazały się fałszywe.

¹¹⁶ Ów łośń nie jest zresztą w nauce jedynym zwierzęciem bohaterem. Zainteresowanym polecam również zapoznanie się z wspomnianym wcześniej badaniem, w którym proszono kurczaki o opinię w kwestii kobiecej urody, by sprawdzić, czy nasze postrzeganie piękna wynika z uwarunkowań biologicznych i ewolucyjnych, czy też z norm kulturowych (Ghirlanda et al. 2002).

¹¹⁷ Oprócz fok, foki są idealne.

¹¹⁸ Jak wynika z badania OECD Better Life Index, o którym pisałam w rozdziale *Czy da się zmierzyć szczęście linijką?*.

¹¹⁹ Superfoods to inaczej superżywność, która ma się odznaczać wysoką wartością odżywczą i prozdrowotną. Nie wszystkie jednak z wymienianych jednym tchem przez ambasadorki żywienia składników mają przebadane właściwości. Przekrojowa metaanaliza dokonana przez dziennikarza danych Davida McCandlessa wskazuje, że wiele z nich może nie mieć żadnego wpływu na nasze zdrowie.

¹²⁰ Jak pokazują badania, nie ma żadnego przekonującego dowodu (poza anegdotycznymi), że protokół Gerstona pomaga w leczeniu raka (National Cancer Institute 2018).

¹²¹ Transmisja społeczna – motywacja do przekazywania sobie informacji (w tym historii). Siłę transmisji społecznej określa właśnie chęć przekazywania historii dalej. Chętnie dzielimy się informacjami, które nas rozbawiły, wzruszyły, wkurzyły, przeraziły, zaskoczyły czy oburzyły moralnie. Dlaczego? Bo chcemy innych rozbawić, wzruszyć, wkurzyć, ostrzec czy zjednoczyć się we wspólnej pogardzie wobec niepożądanego społecznie grupy czy nieakceptowalnych zachowań. Transmisja społeczna nie jest zależna od prawdziwości historii. Przekazujemy sobie historie dające nadzieję bez względu na to, czy są prawdziwe, czy nie. Głównie dlatego, że sami nie mamy prostego instynktu weryfikowania opowieści. Wydają nam się prawdziwe, bo bardzo byśmy chcieli, żeby takie były.

¹²² Zauważcie, że pierwsze litery angielskich odpowiedników układają się w łatwy do zapamiętania akronim STEPPS, czyli takie kroki, tyle że z błędem. Jest coś wspianego w tym, kiedy to natura tak pięknie mówi akronimami.

¹²³ *Pie chart* to wykres kołowy, *donught chart* to taki wykres odwzorowujący irlandzkie rozumienie pączka, w sensie, że taki wykres kołowy z pustą, okrągłą przestrzenią pośrodku.

Opieka redakcyjna: Maciej Makselon
Redakcja: Katarzyna Maciejak
Konsultacja merytoryczna: Maciek Taraday
Korekta: Milena Domańska
Skład i łamanie: Katarzyna Piątek
Projekt okładki: Joanna Strękowska
Ilustracje: Beata Smugaj

Niniejszy produkt objęty jest ochroną prawa autorskiego. Uzyskany dostęp upoważnia wyłącznie do prywatnego użytku osobę, która wykupiła prawo dostępu. Wydawca informuje, że wszelkie udostępnianie osobom trzecim, nieokreślonym adresatom lub w jakikolwiek inny sposób upowszechnianie, upublicznianie, kopiowanie oraz przetwarzanie w technikach cyfrowych lub podobnych – jest nielegalne i podlega właściwym sankcjom.

Ilustracje: Copyright © by Beata Smugaj

Copyright © by Janina Bąk

Copyright © for the Polish edition by Grupa Wydawnicza Foksal, MMXX

Wydanie I

Warszawa MMXX

ISBN: 978-83-280-7260-2

Skład wersji elektronicznej: Michał Olewnik / [Grupa Wydawnicza Foksal Sp. z o.o.](#)
i Michał Latusek / [Virtualo Sp. z o.o.](#)

virtualo